

The international journal of science / 23 January 2020

# nature

## CELLULAR CONNECTIONS

First Asgard archaeon  
cultured in the lab hints  
at how complex  
life evolved

### Coastal view

Global survey reveals  
changing footprint  
of river deltas

### Cancer treatment

How B cells act as  
biomarkers for  
immunotherapy

### 2020 vision

Technology innovations  
that could enhance the  
research toolkit

Vol. 575 No. 7794  
nature.com



## Australia: show the world what climate action looks like

**Scott Morrison's government must act on overwhelming evidence and public opinion.**

**L**ast November, as bush fires began to roar across large swathes of Australia, people started to ask: could such an extreme event be connected to climate change?

Prime Minister Scott Morrison dodged the question. Gladys Berejiklian, the premier of the state of New South Wales, where the fires have had the biggest impact, said that during the unfolding disaster was not the time to talk about climate change. Two months on, this season's devastating conflagrations have killed at least 28 people and an estimated one billion native animals; burnt about 10 million hectares of vegetation; and destroyed more than 2,000 homes.

The top priority is to protect lives and ecosystems. But the nation's leaders must surely realize that they not only need to talk about climate change, but also need to act decisively to reduce the emissions that are driving it.

Australia's leaders have known for many years that climate change would make bush fires worse. They were warned in an independent report commissioned by the national and state governments in 2008 that from 2020 onwards, fire seasons would start earlier, end later and be more intense.

But as *Nature* has frequently reported, the country's politicians delayed meaningful action through a wasted decade of arguments over whether human activities are causing climate change – in the face of overwhelming scientific evidence that they are. Undoubtedly, one reason for this is that Australia – which is the world's largest coal exporter – has repeatedly prioritized the coal industry's needs over the planet's.

### Not enough

The government now says it is on track to reduce greenhouse-gas emissions by 26–28% of 2005 levels by 2030, to meet its commitment under the 2015 Paris climate agreement. Its plan includes a policy to pay farmers and businesses to restore or protect native vegetation, and a programme to encourage energy efficiency.

But commitments on such a scale – whether from Australia or other countries – are insufficient to limit warming to below 2 °C above pre-industrial levels, the goal of the agreement. And a significant portion of Australia's planned cuts is to be achieved through accounting tricks, rather than actual emissions reductions. The government plans for around half – 367 million tonnes of greenhouse



Australian Prime Minister Scott Morrison visiting a fire-hit area in Victoria.

**Australia's leaders have known for many years that climate change would make bush fires worse."**

gases – to come from 'credits' it accumulated by surpassing its targets under the previous climate agreement, the 1997 Kyoto Protocol. That means its actual cuts will be 15% from 2005 levels. No other high-income country that has signed the Paris agreement has said it will transfer its Kyoto credits in this way – and nor should Australia.

Last week, after international outrage over his lack of leadership, Morrison switched gears. He started talking about how, as a result of the catastrophic fires, the government would focus on actions that build resilience and adaptation to extreme events, such as bush fires, heatwaves and droughts.

For Australia, that's a significant move – but it is not enough. The government has to do much more to cut its emissions, too. Just reacting to the impacts of climate change without addressing the cause is like treating people for lung cancer while continuing to let them smoke.

Australia's tragedy is that more-extreme fires are already forecast. Centuries of greenhouse-gas emissions have locked the world into several decades of warming, even if global emissions were to drop to zero now. If the Morrison government continues its current trajectory, then the country is likely to experience even more severe droughts and fires.

The Morrison government has to make a choice: does it want Australians to live with fires that are becoming worse than those in the past but which can still be managed to some extent? Or does it want to put citizens at risk of future fire conditions that are even more catastrophic than this season's? There can be only one answer to this question if the government accepts that its first role is always to protect its citizens and its country.

We frequently hear the argument that actions from individual countries such as Australia will, on their own, make little difference to global warming. But that is why we have global agreements. Change will come when everyone acts in concert. Australia, along with the United States, China, the European Union and others all have to play their part,

JAMES ROSS/GETTY



leading the way on decarbonizing energy for households, industry, transport and more.

Instead of arguing with its climate researchers, Australia's government needs to work with them to accelerate this transition, and to ensure that, as far as possible, lives and livelihoods are protected when change arrives. A country on the front lines of climate change has no other choice.

## Stop the Wuhan coronavirus

**Vigilance, preparedness, speed, transparency and global coordination are now crucial to preventing a new infectious disease from becoming a global emergency.**

**A**s hundreds of millions of people in China take to the roads, railway and skies to be with their families for the new year holidays, authorities in the country and around the world have mounted an enormous operation to track and screen travellers from Wuhan in central China.

This follows the outbreak of a mysterious pneumonia-like coronavirus, first reported on the last day of December 2019, that has so far claimed six lives in China. The World Health Organization is deciding whether to declare the situation an international public-health emergency.

The virus has been spreading. On 21 January, as *Nature* went to press, there were almost 300 reported cases – seven times the figure stated five days earlier. Over the past week, authorities in South Korea, Thailand and Japan have also reported cases. Researchers at Imperial College London who have modelled the outbreak on the basis of estimates of travel out of Wuhan say the virus might have infected as many as 1,700 people.

The virus, which still lacks a formal name, is being called 2019-nCoV. It is a relative of both the deadly severe acute respiratory syndrome (SARS) and the Middle East respiratory syndrome (MERS) viruses. People with the virus report a fever along with other symptoms of lower-respiratory infection such as a cough or breathing difficulties. The first people infected in China are understood to have caught the virus in one of Wuhan's live animal and seafood markets – probably from an animal. Some 95% of the total cases, including those in Japan, South Korea and Thailand, also involved people who had been to Wuhan.

The virus has not been found in humans before and knowledge of how it is spread is still evolving. Last week, government officials and researchers in China who are tracking the virus told *Nature* they didn't think it spreads readily from human to human, at least not as fast as SARS. But this view is being revised following the intervention of SARS specialist Zhong Nanshan. After a visit to Wuhan on

 **Authorities must continue to report what they know and what more they are uncovering."**

20 January, Zhong, who directs the State Key Laboratory of Respiratory Disease in Guangzhou, confirmed that 14 medical workers had been infected by one virus carrier, raising concern that some people might be 'super-spreaders' of the virus. Stopping the further spread of the disease out of Wuhan, possibly by banning infected people from leaving Wuhan, has to be a top priority, he said.

China's health authorities and the government have been moving quickly. Also on 20 January, the national broadcaster reported that president Xi Jinping had ordered that the virus be "resolutely contained", and Premier Li Keqiang announced a steering group to tackle disease spread. At the beginning of the month, local authorities in Wuhan closed and disinfected the animal market, and health authorities have reported the results of their disease surveillance efforts.

Researchers, too, have had a crucial role, in publishing and sharing genome sequences. Four different research groups sequenced the genomes of six virus samples – and analyses of all six agree that the virus is a relative of SARS. Researchers are to be commended for making sequence data available, and they should continue to do so. (Release of such data, as well as deposition of manuscripts on pre-print servers, will not affect the consideration of papers submitted to *Nature*.)

As China's government has recognized, the authorities fumbled in their response to SARS, which spread globally, killing more than 770 people in 2002–03. Fifteen per cent of those infected died, a rate that seems much higher than that of the current outbreak – at least from what is known so far. In contrast to SARS, the response this time has been faster, more assured and more transparent.

But there is still much to do, and quickly. The virus's original source must be confirmed – something that is proving difficult. Researchers have found virus traces in swabs taken from the animal market. The authorities, rightly, made closing and sterilizing the market their first priority, but in their rush to do so they might have missed a chance to test the animals. In the case of SARS, we now know that bats transmitted the virus to other animals, which then passed it to humans. Other questions include confirming the method of transmission for new cases, as well as understanding the virus's ability to cause serious illness. Virus genomes from infected people will need to be sequenced continually to understand the extent to which the virus is evolving.

China's health authorities did well to act more quickly than in the past. Now, they must continue to report what they know and what more they are uncovering. The emerging situation requires global co-ordination and leadership from the World Health Organization, with the support of public-health agencies worldwide. Researchers must work fast, collaboratively and transparently to address the key research questions. The world has had plenty of practice with SARS and avian flu – we should know what to do.

Around 7 million people are preparing to fly from China to 400 cities in 100 countries to celebrate the Chinese New Year. Now is the time to stop this outbreak spiralling into a global health emergency.



# World view

## Our best weapons against cancer are not magic bullets



By Vinay Prasad

**Better health and social policy would save more lives than sophisticated drugs.**

**E**arlier this month, the American Cancer Society announced its latest figures on cancer incidence and mortality (R. L. Siegel. *et al. CA Cancer J. Clin.* 70, 7–30; 2020). These included the largest drop ever observed in national cancer statistics, which several media outlets seized on. Cancer death rates in the United States peaked in 1990, and in 2008–17 fell by about 1.5% per year. Between 2016 and 2017, the drop was slightly larger: 2.2%. This is undeniably good news.

But our optimism must be tempered by other measures of population health – particularly declining life expectancy.

The reason behind the large drop is a decrease in mortality for lung cancer – without lung cancer, the rate is still about 1.5%. Several reactions to the Cancer Society's news heralded advances in precision treatments. Yet much of the continued reduction in mortality is due to the lower incidence of lung cancer, or a reduction in new cases per year. And new drugs cannot cause that. The two major therapeutic advances for treating this cancer – genome-targeted therapies and immunotherapy – are currently approved for the worst-off individuals: those with advanced or metastatic disease.

Exciting technologies that uncover genetic drivers of cancer and unleash the immune system against it make headlines, but I think we must be careful not to give customized treatments too much credit, and I have been outspoken about my work to pin down the impact of these therapies. We would do better to focus on public-health strategies that are less glamorous.

My colleagues and I have estimated that, as of 2018, 8.33% of the US population with advanced cancer was eligible for genome-targeted therapy, up from 5.09% in 2006 (J. Marquart *et al. JAMA Oncol.* 4, 1093–1098; 2018). Another work found that people whose lung cancers are eligible for genome-targeted treatments and who receive them live, overall, about 30 weeks longer than those who are eligible and are not treated (G. Singal *et al. J. Am. Med. Assoc.* 321, 1391–1399; 2019). That benefit is real, but is unlikely to have altered mortality rates markedly across a population.

Similarly, immunotherapy – which expanded into the market in 2015 – might have had only limited effects on the drop in overall cancer mortality. The benefits for melanoma and for advanced and metastatic lung cancer are impressive, but so far affect relatively few people.

Much bigger drops in US cancer mortality would come from a fairer society. The American Cancer

**The data do make it clear that the majority of our most effective solutions will be outside the cabinet of cutting-edge medicines."**

**Vinay Prasad** is associate professor of medicine at Oregon Health & Science University in Portland, and author of the forthcoming book *Malignant: How Bad Policy and Bad Evidence Harm People with Cancer*. e-mail: [prasad@ohsu.edu](mailto:prasad@ohsu.edu) Twitter: [@vprasadmdmph](https://twitter.com/vprasadmdmph)

The author declares competing interests; see [go.nature.com/2tuseqb](https://go.nature.com/2tuseqb) for details.

Society estimates that, in 2014, 59% of lung-cancer deaths observed in people aged 25–74 could have been averted by eliminating socio-economic disparities (R. L. Siegel *et al. CA Cancer J. Clin.* 68, 329–339; 2018).

What's more, US life expectancy has fallen for three straight years. The cause is largely diseases of despair: drug overdose, suicide and alcohol-related liver disease. And these kinds of risk factor cluster. People who die from using opiates are more likely to smoke, for instance. The American Cancer Society uses age-standardized populations to address concerns that a rise in untimely deaths could mask what would have been future cancer deaths and thus spuriously improve cancer death statistics, but it is hard to know exactly how factors behind declining life expectancy play into cancer mortality.

The data do make it clear that the majority of our most effective solutions will be found outside the cabinet of cutting-edge medicines. If we want to do all that we can to reduce the burden of cancer and to improve life expectancy, we must harness the tools of population statistics.

That means we need to create strategies to treat hypertension, end the use of tobacco products, dismantle policies that promote obesity and use of environmental carcinogens, encourage physical activity and reduce levels of carcinogens in the environment. In my cancer clinic, I often wish I had more effective drugs for the person in front of me. I, too, want sophisticated treatments that work. But what I really wish is that the person I'm treating did not have cancer at all.

Our public policy is a series of self-inflicted wounds. The current US administration has allowed loopholes that let the known carcinogen asbestos remain in use. It has failed to improve standards for airborne particulate pollution, clearly linked to higher rates of diseases and death. It reversed a decision to ban a pesticide, chlorpyrifos, associated with impaired childhood brain development, and atrazine, linked to leukaemia.

My deep frustration is this: it is hard to escape the conclusion that we, as a society, are not doing what it takes to maximize our health. We are prioritizing medications that cost US\$100,000 a year or more, and at the same time are loosening restrictions on environmental pollution. These policies have one thing in common: they enhance corporate profits. It will take a realignment of public policy to make sure that we pursue systems that instead prioritize health.

Public-health policies are not personalized to any individual, but can promote longevity for all of us, even if it will not make for feel-good stories about scientific breakthroughs or miraculous drugs. In this exciting age of precision medicine, we will reap the biggest gains by celebrating better health for everyone.



# News in brief

## CHINA NEARS TOP SPOT FOR RESEARCH SPENDING

The gap in research funding between the United States and China is closing fast, despite modest increases in US funding since 2000, according to statistics assembled by the US National Science Foundation (NSF).

From 2000 to 2017, research and development (R&D) spending in the United States grew at an average of 4.3% per year, the NSF found. But spending in China grew by more than 17% per year during the same period. Several other countries, including Germany and South Korea, also increased their spending at rates that outstripped that of the United States, although they remain solidly behind the two global leaders in terms of total funding. The United States accounted for 25% of the US\$2.2 trillion spent on R&D worldwide in 2017, with China making up 23%.

The figures come from the latest edition of the NSF's biennial *Science and Engineering Indicators* report, which compiles metrics on the state of science and engineering in the country. The United States is increasingly "seen globally as an

important leader rather than the uncontested leader" in science and engineering, according to the report, released on 15 January.

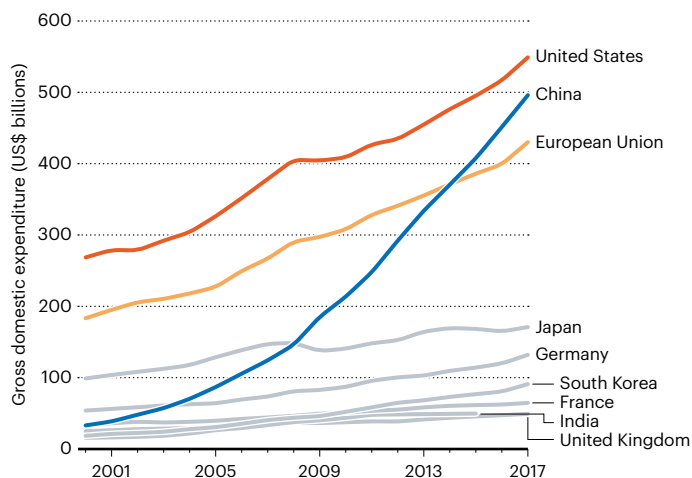
Preliminary data from 2019 suggest that China has already surpassed the United States in R&D spending, said Julia Philips, chair of the National Science Board's science and engineering policy committee, during a press briefing. The board oversees the NSF and produces the *Indicators* reports.

The emergence of innovation powerhouses outside the United States "can only be good", says Diane Souvaine, a computer scientist at Tufts University in Medford, Massachusetts, who chairs the National Science Board. She notes that the United States still leads the world in many important metrics, such as total investment in R&D, proportion of highly cited publications and enrolment of internationally mobile students.

However, the NSF report found that the number of foreign-born students enrolling in US universities has declined slightly in recent years.

## SCIENCE SPENDING

China is catching up to the United States on funding for research and development.



## NEW VIRUS SURGING IN ASIA CAUSES ALARM

Scientists are increasingly concerned about a new virus that is spreading in Asia. As *Nature* went to press, Chinese officials had reported 291 cases nationwide, most in the city of Wuhan, where the outbreak began. Thailand, Japan and South Korea are among the nations that have reported infections. At least six people have died from the virus, which causes a respiratory illness.

Chinese officials have also confirmed that the virus can spread from person to person, although the extent of such transmissibility is unclear. The surge in infections is alarming because of Chinese New Year this weekend, when hundreds of millions of people will travel to their home towns or overseas. "This could be the beginning of a disaster," says Seungtak Kim, a virologist at the Pasteur Institute Korea in Seongnam, South Korea.

The illness was first detected last December among people who had visited a live-animal market in Wuhan. Scientists have identified the pathogen as a coronavirus, from the same family that causes severe acute respiratory syndrome, or SARS. As *Nature* went to press, the World Health Organization was set to meet on 22 January to decide whether to declare a public-health emergency over the virus.



## High risk of major eruption at Taal volcano





Researchers in the Philippines are monitoring the Taal volcano closely for signs of a major eruption. The volcano's activity has eased since it began spewing steam and ash more than a week ago, but the threat of a large-scale eruption remains, say scientists. In addition to the immediate risk to life, such an event could contaminate water supplies and disrupt power generation for millions, and halt ground and air travel.

At 2.30 p.m. local time on 12 January, Taal started ejecting lava and blew out a giant plume of rock fragments. Ash travelled as far north as Quezon City, some 70 kilometres away, forcing tens of thousands of people on Taal's Volcano Island and in nearby provinces to flee.

The volcano's activity has stalled, but this does not mean the worst is over, says Mariton Bornas, a volcanologist at the Philippine Institute of Volcanology and Seismology just north of Manila.

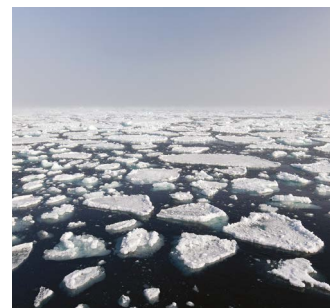
The volcano remains at level 4 on the country's volcano-alert system, the second-highest level, meaning a hazardous eruption could happen in hours or days.

## CATASTROPHIC AUSTRALIAN BUSH FIRES DERAIL RESEARCH

The blazes raging across Australia have damaged lives, homes and businesses. They have also destroyed scientific equipment and derailed research.

Remote-sensing specialist Will Woodgate at the University of Queensland in Brisbane manages a site in the Bago State Forest that gathers data on land surface conditions to feed into global climate models. As fire tore through the site on New Year's Eve, the data that have flowed from it for 20 years stopped. Photos suggest that the layer of vegetation under the forest canopy has been wiped out, although the canopy itself is intact. Woodgate says sensors at the top of a tower at the centre of the site could have survived.

Elsewhere, the Australian Mountain Research Facility was set up last year by the Australian National University in Canberra to study how a changing climate affects alpine landscapes. It had planned to deploy sensors and monitoring equipment to its eight field sites in the Australian summer. But fire at one site has left "nothing but bare soil", says soil scientist Zach Brown, the senior technical officer for the project. Installation of equipment across the network has been set back by a year, he says.



## OZONE-EATING GASES LINKED TO EXTREME ARCTIC WARMING

Gases that deplete Earth's protective ozone layer could be responsible for up to half of the effects of climate change observed in the Arctic from 1955 to 2005.

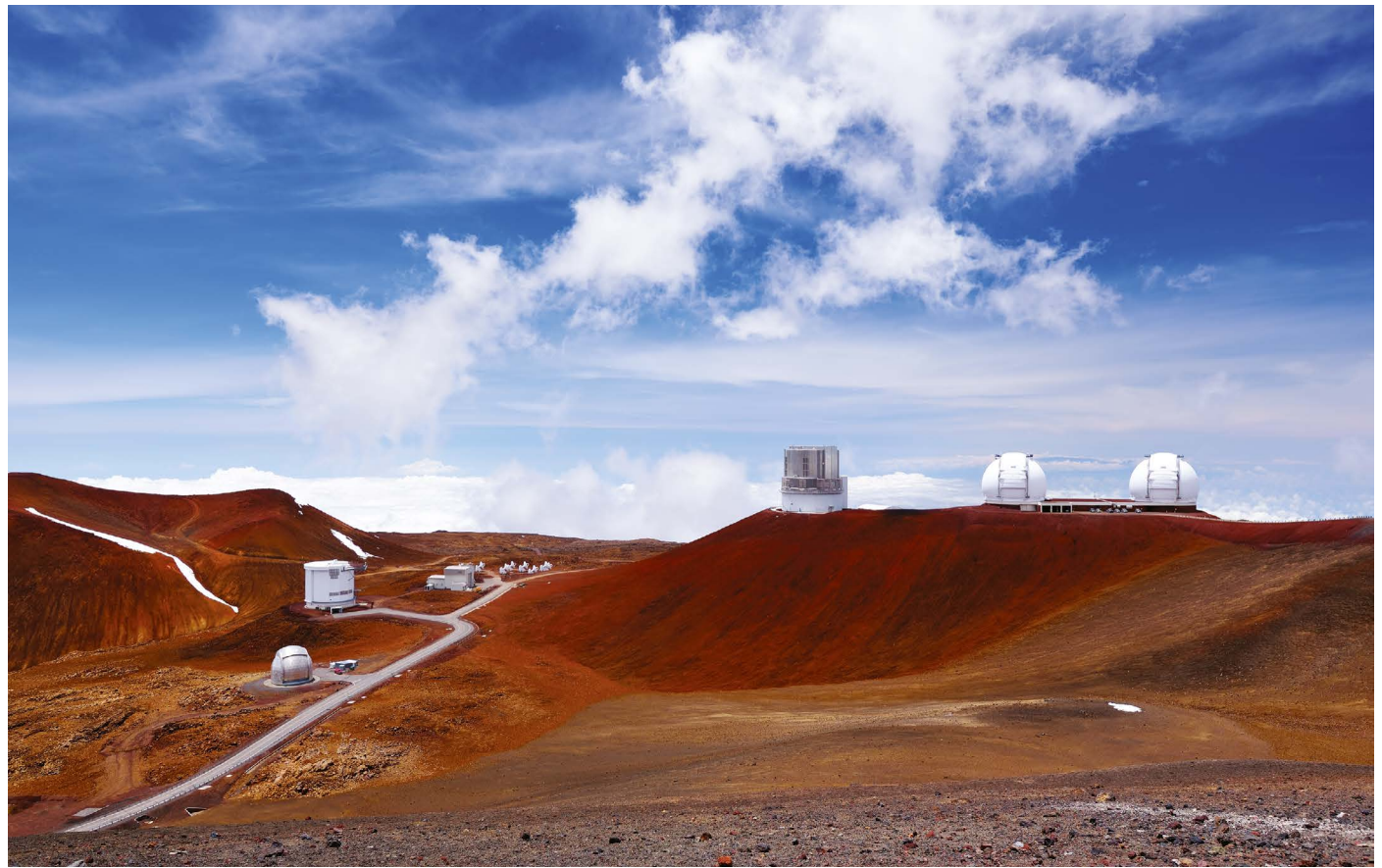
The finding, published on 20 January, could help to explain the disproportionate toll that climate change has taken on the region, an effect that has long puzzled scientists (L. M. Polvani *et al. Nature Clim. Change* [http://doi.org/djt5](https://doi.org/djt5); 2020). The Arctic is warming at more than twice the average rate of the rest of the globe – a phenomenon known as Arctic amplification – and it is losing sea ice at a staggering pace.

Ozone-depleting substances, including chlorofluorocarbons (CFCs), are known to heat the atmosphere more efficiently than carbon dioxide. But most research on these chemicals has focused on their effects on the ozone layer.

A team of researchers compared climate simulations both with and without the mass emission of CFCs that began in the 1950s. Without CFCs, the simulations showed an average Arctic warming of 0.82 °C, but with CFCs, the number jumped to 1.59 °C.

Replicating these results in multiple climate models will be crucial for improving estimates of how much responsibility CFCs bear for heating the Arctic, say researchers.

# News in focus



The summit of Mauna Kea in Hawaii already hosts 13 telescopes.

## HOW THE FIGHT OVER A HAWAII MEGA-TELESCOPE COULD CHANGE ASTRONOMY

Thirty Meter Telescope controversy is forcing scientists to grapple with how their research affects Indigenous peoples.

By Alexandra Witze  
on Mauna Kea, Hawaii

One morning earlier this month, on the rain-soaked slopes of Mauna Kea in Hawaii, Noe Noe Wong-Wilson was settled in for the long haul. Wrapped in a trench coat to keep out the wind and cold, the educator and activist held a meeting amid camp beds and folding chairs inside a giant tarpaulin-sheltered tent.

Wong-Wilson is a leader of the Mauna Kea *kia'i*, a group of Native Hawaiians who have been encamped near the volcano's base since

last July. They are preventing construction workers from building an enormous telescope near the summit, on land the *kia'i* regard as sacred. The planned Thirty Meter Telescope (TMT) would transform astronomy by peering into the Universe with sharper vision than that of nearly any other. But there are already 13 telescopes atop Mauna Kea, and the *kia'i* say that adding the TMT would be too much.

If project officials cannot work out a way to build the telescope in Hawaii, they intend to move it to an alternative – but slightly less scientifically compelling – site in Spain's Canary Islands. Whatever the outcome,

the debate over the TMT is profoundly transforming how astronomy is done in Hawaii. The island chain – one of the world's best places for stargazing – has become a testing ground for the ethics of conducting research in a place full of injustice towards Indigenous peoples.

“Gone are the days of the scientific conceit of being separate from the community,” says Jessica Dempsey, deputy director of the East Asian Observatory, which operates a telescope on Mauna Kea. “Astronomers really have to do more contemplation about where they are in the world, and about the social context and



impact of their work.”

How the Mauna Kea stand-off plays out could affect astronomical research in other locations and other fields of science around the world, she says.

Astronomers confronted this new reality this month, when thousands of them attended a meeting of the American Astronomical Society in Honolulu. The conference featured many sessions on Hawaiian culture and astronomy and saw anti- and pro-TMT demonstrations. “It’s an industry that is congruent with our culture as explorers,” said Malia Martin, a Native Hawaiian who supports the TMT, as she waved a Hawaiian flag outside the convention centre.

### Changing course

The fight over the TMT has become a symbol of historical inequities in Hawaii, notably the seizure of lands from Native Hawaiians before and after the United States annexed the islands in 1898. “This is a political issue rooted in historical injustice,” says Greg Chun, executive director of Mauna Kea Stewardship for the University of Hawaii, which manages the mountaintop land on which the observatories sit. Homes and vehicles across the islands often fly the Hawaiian flag upside down as a symbol of protest against the US government.

TMT officials have tried to address some of these long-standing issues, in part by establishing educational and workforce-training programmes for local residents. But the project, which is expected to cost its partners in the United States, India, China, Japan and Canada more than US\$1.4 billion, has not been able to proceed with construction. Both times it tried – first in 2015, and then again in July 2019 – the *kia’i* blocked the road to Mauna Kea’s summit.

The 13 existing telescopes atop the mountain face an uncertain future. The University of Hawaii has committed to removing five as a condition of the permit to build the TMT. The three chosen so far are among the oldest telescopes on Mauna Kea.

The future of the rest – which include some of the world’s most scientifically productive observatories, such as the Keck and Canada-France-Hawaii telescopes – is assured only until 2033. Astronomy will end on Mauna Kea after that if the state government does not renew the university’s master lease on the mountaintop, which governs all the telescopes’ operations.

From her spot at the base of the mountain, Wong-Wilson says she is open to the possibility of the lease being renewed. “There is space for discussion about improving the way astronomy remains upon our mountain,” she says. “But attitudes have to change. Astronomers look at us like we’re the bad guys, like we’re intruding on their space. It’s quite the opposite: they’re in our space.”

Cutting-edge astronomy should continue within the footprint of the existing observatories, says Rosie Alegado, an oceanographer at the University of Hawaii at Manoa. She helps lead a group of Native Hawaiian scientists who this month called for an immediate halt to the TMT project while organizers seek “informed consent” for the telescope to move forward

**“Gone are the days of the scientific conceit of being separate from the community.”**

(S. Kahanamoku *et al.* Preprint at <https://arxiv.org/abs/2001.00970>; 2020). They also called for Indigenous people to have more overall input into decisions involving the mountain. “I feel like astronomy on Mauna Kea could represent an example of when science got off course, but we course-corrected and came back stronger than ever,” she says.

### Momentous decision

How that might happen remains to be seen. If the TMT moves to the Canary Islands, it will take with it money it would otherwise spend to help maintain the infrastructure for

astronomy on Mauna Kea, such as the road to the summit. The move could also shift the focus of TMT partners, a few of whom operate some of the existing telescopes, away from Hawaii.

State and local governments have brokered a detente between TMT officials and the *kia’i* until the end of February. Representatives of various groups are meeting to try to hammer out some sort of agreement for whether and how the TMT might proceed on Mauna Kea.

But the clock is ticking. The telescope needs funding from the US National Science Foundation to keep moving forward. To get it, the project would need to be ranked highly in the next ‘decadal’ survey of priorities for US astronomy, which scientists are compiling. Results are expected in early 2021. The TMT might not get a high ranking if it can’t show a clear path to construction – which means that the issues with Mauna Kea need to be sorted out, or it needs to move to the Canaries.

For Dempsey, the debate has pushed long-simmering disagreements over science and land rights to the fore. “I’m kind of glad in some ways that we’ve been forced into this conversation,” she says. “We didn’t do enough creative things in our local community in Hawaii until we were forced to – by people saying that this is not okay.”

## SUPERCOMPUTER SCOURS FOSSIL RECORD FOR HIDDEN EXTINCTIONS

Palaeontologists have charted 300 million years of Earth’s history in breathtaking detail.

By Ewen Callaway

**P**alaeontologists have a fuzzy view of Earth’s history. An incomplete fossil record and imprecise dating techniques make it hard to pinpoint events that happened within geological eras spanning millions of years. Now, a period that saw a boom in animal complexity and one of Earth’s greatest mass extinctions is coming into sharp focus.

Using the world’s fourth most powerful supercomputer, Tianhe II, a team of scientists based mostly in China mined a fossil database of more than 11,000 species that lived during the period from around 540 million to 250 million years ago. The result is a history of life during this period, the early Palaeozoic era, that can pinpoint the rise and fall of species during diversifications and mass extinctions

to within about 26,000 years (J.-x. Fan *et al. Science* **367**, 272–277; 2020).

“It is kind of amazing,” says Peter Wagner, a palaeontologist and evolutionary biologist at the University of Nebraska–Lincoln, who was not involved in the work. Being able to look at species diversity on this scale is like going from a system where “people who lived in the same century are considered to be contemporaries, to one in which only people who lived during the same 6-month period are deemed to be contemporaries”, he wrote in an essay accompanying the study (P. Wagner *Science* **367**, 249; 2020).

Such a view, Wagner adds, will help scientists to identify the causes of mass extinctions – such as the event at the end of the Permian period, some 252 million years ago, that wiped out more than 95% of marine species – as well as understand less dramatic species die-offs



Trilobites disappeared from the fossil record during a mass extinction 252 million years ago.

and rebounds that have been hard to uncover because of gaps in the fossil record. An understanding of these processes could reveal parallels to the planet's current loss of biodiversity.

### Patchy record

Most organisms in Earth's history didn't leave fossils, and scientists have identified only a tiny fraction of those that did. As a result, it can be hard to tell whether changes in the fossil record mark real shifts, such as mass extinctions, or are simply caused by a lack of fossil finds.

In the 1960s, palaeontologists began analysing the fossil record systematically, revealing multiple mass extinctions and periods during which life flourished. But these and later efforts could usually pinpoint biodiversity changes to within only about ten million years, because fossils were lumped into relatively long geological periods and analysed en masse.

To improve on this, a team led by palaeontologist Jun-xuan Fan at Nanjing University in China created and analysed a database of fossil marine invertebrates that were found in more than 3,000 layers of rock, mostly from China but representing geology around the planet during the early Palaeozoic. The group then used software to measure when individual species had emerged and gone extinct.

The program took advantage of the fact that species were usually found in multiple rock formations – each spanning hundreds of thousands to millions of years – and used this information to place upper and lower limits on the period in which the species actually existed. The effort revealed for how long, and in what order, all 11,000 species had existed. It took the supercomputer around seven million processor hours.

Using this approach, the team was able to learn extra details about events such as the end-Permian extinction, and the Cambrian explosion around 540 million years ago. The analysis showed, for instance, that species diversity declined in the 80,000 years leading up to the end-Permian mass extinction, which itself occurred over about 60,000 years.

The findings also cast doubt on the existence of a smaller-scale die-off known as the end-Guadalupian extinction, which is thought to have wiped out many marine species around 260 million years ago. That was the biggest surprise, says Mike Benton, a palaeontologist at the University of Bristol, UK, who has documented changes in vertebrate diversity during that period. The study, he adds, “represents a pretty amazing big-data endeavour”.

Benton hopes to see the effort extended to later periods – particularly the past 100 million years. Palaeontologists disagree over whether an apparent increase in animal diversity in this period is the result of sampling bias.

Norman MacLeod, a palaeontologist at the University of Nanjing and a co-author of the study, says the team's work might help to reveal the underlying causes of changes in biodiversity, by charting ups and downs on a timescale that can be matched with environmental and climatic shifts.

Wagner adds that the team's approach will be most valuable in uncovering – and explaining – smaller-scale extinctions, not dissimilar to those occurring today. Such extinctions could turn out to be “a bad 100,000 years, or a bad week” for some groups of organisms but not others, he says. “When you get this resolution, it starts opening the doors to actually testing what the smaller-turnover events might be like.”

## STUDIES OF EMBRYO-LIKE STRUCTURES STRUGGLE TO WIN US GRANTS

Biologists say they need clearer guidelines on funding rules for this nascent field.

By Nidhi Subbaraman

Scientists can now create clumps of cells that resemble human embryos, raising hopes that they could study the elusive first stages of human development while avoiding the ethical concerns that make it difficult to study actual human embryos. But as these embryo models – in which human stem cells are transformed into embryo-like structures whose growth mirrors stages of embryonic development – grow in popularity, US researchers say that they are finding it increasingly difficult to obtain federal funding for such work.

The US National Institutes of Health (NIH) in Bethesda, Maryland, has funded and still does fund work on embryo-like structures.

A spokesperson told *Nature* that the agency considers grant applications involving models that “could be considered an organism” on a “case-by-case basis”, and cited a provision of federal law known as the Dickey-Wicker Amendment, which bars the government from funding research that creates or destroys human embryos.

But the ban, which dates back to 1996, was put in place before the advent of techniques that produce embryo-like structures from stem cells. Scientists working on such research say that they need clearer guidance on what is eligible for federal funding. “The writing on the wall is that synthetic embryos are out of bounds with the NIH. The next step in the science is not allowed,” says Eric Siggia, a physicist who studies developmental systems



## News in focus

at the Rockefeller University in New York City.

Amid this growing criticism, the agency's Office of Science Policy asked the US National Academies of Sciences, Engineering, and Medicine (NASEM) to host a day-long workshop to lay out the latest developments in experiments with embryo-like structures. At the NIH's request, the meeting on 17 January in Washington DC did not include any presentations on ethics or regulations.

The NASEM meeting was intended to help people to "better understand some of the unknowns associated with this nascent field", Carrie Wolinetz, the NIH's acting chief of staff and associate director for science policy, wrote in a blogpost last year. "Can research involving various models of aspects of human embryo development be supported by NIH? The answer is 'it depends'," she added.

### Sticky wicket

Embryo research in the United States has long been fraught. In addition to the Dickey-Wicker Amendment, US scientists are guided by an internationally acknowledged ethical guideline called the 14-day rule. This limits embryo research to the two-week period after fertilization. And last June, the US government halted fetal-tissue research by government scientists and began requiring that any grant application involving such material undergo an extra ethics review.

None of these laws and guidelines specifically deals with the increasingly complex collections of cells that mimic the early stages of human embryonic development, and can shed light on processes that are otherwise difficult to study. Crucially, embryo-like structures are not formed from an egg and sperm, as real embryos are. Scientists say that it is unclear whether or how existing guidelines are being applied to research that uses the structures.

Siggia and a colleague at Rockefeller, developmental biologist Ali Brivanlou, submitted a progress report to the NIH in 2018 on their grant to study the mechanisms by which colonies of embryonic stem cells organize themselves. Siggia says that they were told by NIH staff to cut plans for research in which synthetic embryonic cells would interact with "extra-embryonic" cells – tissue that grows into the placenta and other structures that nourish an embryo. "The mix of extra-embryonic and embryonic cells could get what someone would construe to be an embryo – and they didn't want to go anywhere near that," Siggia says. But he argues that the work would be the next logical step in experimental design.

He and Brivanlou resubmitted their plans for the next year after altering the original text. "Then it moved forward," he says.

The Rockefeller group is not the only one adapting its plans so that it can continue its work. Aryeh Warmflash, a stem-cell biologist at Rice University in Houston, Texas, says he

isn't applying for federal funding for work that uses embryo-like structures to study the phase of development known as gastrulation. "It doesn't seem to me to be worth the effort," Warmflash says. He is turning to private funders.

And Fu Jianping, a bioengineer at the University of Michigan in Ann Arbor, says that he submitted a grant application to the NIH to study the origin of cells that are precursors to eggs and sperm using embryo-like structures. The agency reviewed and scored it last June, and a programme officer e-mailed Fu a list of questions, including one that asked whether his experiments would involve extra-embryonic tissue. Several months later, Fu says he hasn't received any funding. "The uncertainty from the funding agencies is definitely going to be a roadblock to continued progress," he says.

An NIH spokesperson told *Nature* that scientists with questions about any grant application or award could contact the relevant agency official, and that the agency does not comment on unfunded grant applications.

The International Society for Stem Cell Research in Skokie, Illinois, said on 16 January that it would release updated guidelines in early 2021 to address the complexity of research with embryo-like structures. It also released a series of recommendations for researchers to follow until then.

"The NIH of course is struggling with the question when is an embryo not an embryo," says Janet Rossant, a developmental biologist at the Hospital for Sick Children in Toronto, Canada, and an organizer of the NASEM workshop. "I would also absolutely say we're not close to a line that should not be crossed."

## HUGE SURVEY REVEALS PRESSURES OF SCIENTISTS' LIVES

### Global study highlights long hours, poor job security and mental-health struggles.

By Alison Abbott

A survey of more than 4,000 scientists has painted a damning picture of the culture in which they work, suggesting that highly competitive and often hostile environments are damaging the quality of research.

Around 80% of the survey's participants – mostly academic researchers in the United Kingdom – believed that competition had fostered mean or aggressive working conditions, and half described struggles with depression or anxiety. Nearly two-thirds of respondents reported witnessing bullying or harassment and 43% said they had experienced it.

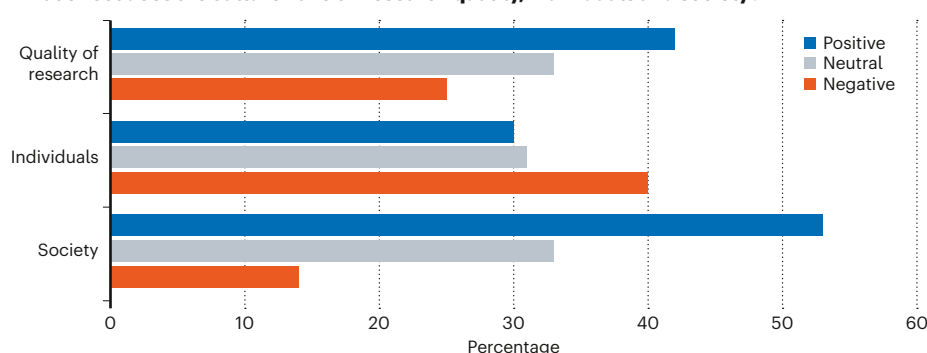
### COST OF THE CULTURE

In a global survey of around 4,000 researchers, 55% said that they had a negative impression of scientific working cultures. One-quarter said that the culture damaged the quality of research.

#### How would you describe research culture?



#### What effect does the culture have on research quality, individuals and society?



SOURCE: WHAT RESEARCHERS THINK ABOUT THE CULTURE THEY WORK IN (WELLCOME, 2020)

“These results paint a shocking portrait of the research environment – and one we must all help change,” says Jeremy Farrar, director of Wellcome, a major research funder in London that conducted the study with market-research agency Shift Learning. “A poor research culture ultimately leads to poor research.”

Farrar says that Wellcome – which supports some 15,000 people working in science worldwide – is committed to addressing the issues highlighted by the survey, and he calls on the entire research system to get on board. “The pressures of working in research must be recognized and acted upon by all, from funders to leaders of research and to heads of universities and institutions,” he says.

### Unsustainable environment

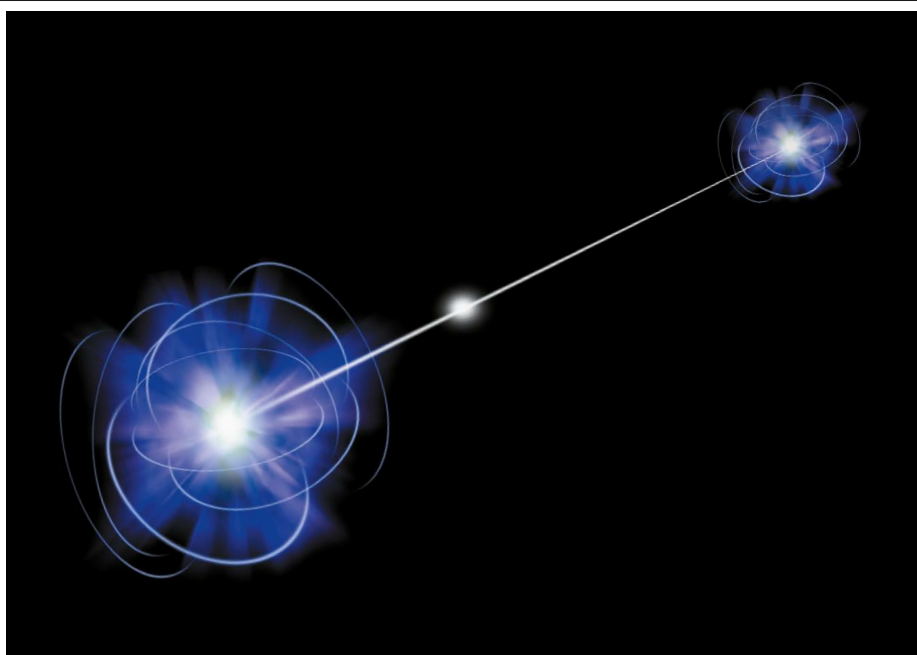
Wellcome conducted the survey, published on 15 January, as part of a broader drive to improve working environments in science. It says the push for excellence has created a troubling culture. “It’s more than clear that our current research practice is not sustainable,” says Beth Thompson, who leads Wellcome’s research-culture initiatives. “We knew things were not right, from our own discussions with scientists, from high-profile bullying cases, reports of misconduct and irreproducibility.”

The results come from an online survey open to all researchers, which was answered by around 4,300 people across career stages and disciplines. Respondents hailed from 87 countries; three-quarters were in the United Kingdom. Workshops with 36 UK-based researchers and in-depth interviews with 94 also informed the findings.

Most researchers reported having pride in their institutions and passion for their work, but spoke of the high personal toll of their environment (see ‘Cost of the culture’). Many accepted that pressure and long hours came with the territory – two-thirds of respondents said they worked for more than 40 hours a week. But researchers said that the situation was worsening and that the negative aspects were no longer offset by job security and the ability to work autonomously, flexibly and creatively. Barely 30% of respondents felt that there was job security in research careers.

Many blamed funders and institutes that emphasize performance indicators and metrics such as number of publications and the impact factors of journals in which researchers publish. They said that the importance of these metrics is often stressed in ways that reduce morale and encourage researchers to game the system. Some said that good management could shelter scientists from such distorting pressures, but that it was too seldom applied.

One-quarter of respondents thought that the quality of research suffered in the unsupportive environments. The same proportion had felt pressured by their supervisors to produce a particular result.



Quantum entanglement is at the centre of a new mathematical proof.

## THE ‘SPOOKINESS’ OF QUANTUM PHYSICS COULD BE INCALCULABLE

Proof at the nexus of pure mathematics and algorithms puts ‘quantum weirdness’ on a new level.

By Davide Castelvecchi

**A**lbert Einstein famously said that quantum mechanics should allow two objects to affect each other’s behaviour instantly across vast distances, something he dubbed “spooky action at a distance”<sup>1</sup>. Decades after his death, experiments confirmed this. But, to this day, it remains unclear exactly how much coordination nature allows between distant objects. Now, five researchers say that they have solved a theoretical problem that shows that the answer is, in principle, unknowable.

The team’s proof<sup>2</sup>, presented in a 165-page paper, was posted on the arXiv preprint repository on 14 January, and has yet to be peer reviewed. If it holds up, it will solve in one fell swoop a number of related problems in pure mathematics, quantum mechanics and a branch of computer science known as complexity theory. In particular, it will answer a mathematical question that has gone unsolved for more than 40 years.

If their proof checks out, “it’s a super-beautiful result” says Stephanie Wehner, a theoretical quantum physicist at Delft University of

Technology in the Netherlands.

At the heart of the paper is proof of a theorem in complexity theory, which is concerned with efficiency of algorithms. Earlier studies had shown this problem to be mathematically equivalent to the question of spooky action at a distance – also known as quantum entanglement<sup>3</sup>.

### Quantum game theory

The theorem concerns a game-theory problem, with a team of two players who are able to coordinate their actions through quantum entanglement, even though they are not allowed to talk to each other. This allows both players to ‘win’ much more often than they would without quantum entanglement. But it is intrinsically impossible for the two players to calculate an optimal strategy, the authors show. This implies that it is impossible to calculate how much coordination they could theoretically achieve. “There is no algorithm that is going to tell you what is the maximal violation you can get in quantum mechanics,” says co-author Thomas Vidick at the California Institute of Technology in Pasadena.

“What’s amazing is that quantum



## News in focus

complexity theory has been the key to the proof,” says Toby Cubitt, a quantum-information theorist at University College London.

News of the paper spread quickly through social media after the work was posted, sparking excitement. “I thought it would turn out to be one of those complexity-theory questions that might take 100 years to answer,” tweeted Joseph Fitzsimons, chief executive of Horizon Quantum Computing, a start-up company in Singapore.

“I’m shitting bricks here,” commented another physicist, Mateus Araújo at the Austrian Academy of Sciences in Vienna. “I never thought I’d see this problem being solved in my lifetime.”

### Observable properties

On the pure-maths side, the problem was known as the Connes’ embedding problem, after the French mathematician and Fields medalist Alain Connes. It is a question in the theory of operators, a branch of maths that itself arose from efforts to provide the foundations of quantum mechanics in the 1930s. Operators are matrices of numbers that can have either a finite or an infinite number of rows and columns. They have a crucial role in quantum theory, whereby each operator encodes an observable property of a physical object.

In a 1976 paper<sup>4</sup>, using the language of operators, Connes asked whether quantum systems with infinitely many measurable variables could be approximated by simpler systems that have a finite number.

But the paper by Vidick and his collaborators shows that the answer is no – there are, in principle, quantum systems that cannot be approximated by ‘finite’ ones. According

**“I thought it would turn out to be one of those questions that might take 100 years to answer.”**

to work by physicist Boris Tsirelson<sup>5</sup>, who reformulated the problem, this also means that it is impossible to calculate the amount of correlation that two such systems can display across space when entangled.

### Disparate fields

The proof has come as a surprise to much of the community. “I was sure that Tsirelson’s problem had a positive answer,” commented Araújo on one blog, adding that the result shook his basic conviction that “nature is in some vague sense fundamentally finite”.

But researchers have barely begun to grasp

the implications of the results. Quantum entanglement is at the heart of the nascent fields of quantum computing and quantum communications, and could be used as the basis of super-secure networks. In particular, measuring the amount of correlation between entangled objects in a communication system can provide proof that it is safe from eavesdropping. But the results probably do not have technological implications, Wehner says, because all applications use quantum systems that are finite. In fact, it could be difficult to even conceive an experiment that could test quantum weirdness on an intrinsically infinite system, she says.

The confluence of complexity theory, quantum information and mathematics means that there are very few researchers who say that they are able to grasp all the facets of this paper. Connes himself told *Nature* that he was not qualified to comment. But he added that he was surprised by how many ramifications it has turned out to have. “It is amazing that the problem went so deep and I never foresaw that!”

1. Einstein, A., Podolsky, B. & Rosen, N. *Phys. Rev.* **47**, 777 (1935).
2. Ji, Z., Natarajan, A., Vidick, T., Wright, J. & Yuen, H. <https://arxiv.org/abs/2001.04383> (2020).
3. Vidick, T. et al. *Not. Am. Math. Soc.* **66**, 1618–1627 (2019).
4. Connes, A. *Ann. Math.* **104**, 73–115 (1976).
5. Tsirelson, B. *Hadronic J. Suppl.* **8**, 329–345 (1993).



**The week’s best science,  
from the world’s leading  
science journal.**

**NATURE.COM/NATURE/PODCAST**

**nature**

# The pollution detectives

Someone, somewhere, is producing banned ozone-destroying chemicals. Meet the researchers tracking down the rogue polluters who are putting the planet at risk. **By Jane Palmer**

**H**igh in the Swiss Alps, scientists in a small research station are busy fingerprinting the atmosphere.

Perched on a mountain ridge at around 3,450 metres altitude, the Jungfraujoch centre boasts five laboratories, a workshop, a library, a tiny kitchen and ten small bedrooms. Day and night, funnels suck in the thin mountain air and channel it into a series of instruments designed to separate, identify and measure the chemicals swirling through this pristine locale. “We are scanning the whole spectrum of thousands and thousands of molecules,” says atmospheric chemist Martin Vollmer. “It is like we are taking the DNA of the atmosphere.”

Vollmer, who works at the Swiss Federal Laboratories for Materials Science and Technology (EMPA) in Dübendorf, specializes in sniffing out newly emerging trace gases, which make up less than 1% by volume of the planet’s atmosphere. Some of the most notorious are the chlorofluorocarbon (CFC) coolants used for refrigeration and foam production. These destroy the ozone layer, the shield that protects life on Earth from damaging ultraviolet light. In 1987, after researchers demonstrated the threat posed by CFCs, nations banded together to adopt an international agreement known as the Montreal Protocol, to control and eventually phase out CFCs. Updates to the treaty have outlawed some of their replacements, which also turned out to damage the ozone layer, climate or both.

Behind the scenes, scientists such as Vollmer are keeping watch over the health of the atmosphere – in part to make sure nations are honouring their promises. “This is detective work,” says Stephen Montzka of the US National Oceanic and Atmospheric Administration (NOAA) in Boulder, Colorado. “Our remit is to understand if things are changing as expected.”

For many years, the news coming from these air-monitoring campaigns was good. Concentrations of CFCs and several other dangerous compounds were declining steadily. It was the biggest win in environmental policy the world has ever seen, say researchers.

Then, in May 2018, Montzka reported a disturbing blip: levels of one of the most harmful chemicals, trichlorofluoromethane, known as CFC-11, weren’t dropping as fast as expected<sup>1</sup>, suggesting that companies were producing this compound somewhere, in violation of the protocol. “It was the most surprising and shocking thing I’ve seen in my entire career,” Montzka says.

Montzka’s research pointed to eastern Asia, and a follow-up study last May pinpointed the source of a significant fraction of the emissions to two provinces in China<sup>2</sup>. The discovery of these rogue CFC-11 emissions has highlighted just how much the Montreal Protocol relies on the vigilance of scientists. But it has also raised questions about whether researchers can keep up with an ever-growing list of damaging compounds – some so new that their impacts remain unknown.

For the moment, they hope they are winning. Last November, nations that are parties to the Montreal Protocol gathered in Rome, where Montzka presented some positive news about the illegal CFC emissions.

## Fresh start

It all starts with fresh air. Every week, come rain, shine or, more typically, snow, Jen Morse makes the trek up to a small green shack on Colorado’s Niwot Ridge, which lies on the Front Range of the southern Rocky Mountains. In summer, she can drive part of the way and has to hike only the final kilometre of the 6-kilometre trip; in winter, she has to ski the entire distance to the remote, wind-swept spot at 3,523 metres altitude, carrying four large gas canisters in her backpack.

Once in the shack, Morse, who is a climate technician at the University of Colorado, Boulder, connects each flask to an inlet and waits for them to fill. She then heads back down and delivers the snapshots of mountain air to NOAA’s Global Monitoring Division in Boulder, just 40 kilometres away. At the lab, Montzka and his colleagues run the flasks’ contents through three separate gas chromatographs to determine what resides in the ‘background’ atmosphere, which doesn’t have any nearby contamination and therefore provides a reading of chemicals circling the entire globe. “We have to pick special locations far away from local sources of pollution to do that,” Montzka says. “These are desolate areas that are hard and expensive and difficult to be at.”

Flasks are shipped to the lab from 16 sites around the world, including the South Pole, the top of Greenland’s ice cap and the tip of Tasmania in Australia.

The NOAA team runs samples through its instruments to determine the levels of 50 trace gases in the atmosphere. The Jungfraujoch lab is part of a second, NASA-sponsored network called the Advanced Global Atmospheric Gases Experiment (AGAGE), which has 13 active stations in a dozen nations.

Some of these sites have been monitoring CFCs and related compounds since the 1970s. When these compounds were invented in the 1920s, chemists regarded them as safe. But by the 1970s, researchers recognized that CFCs could drift up to the stratosphere and erode the protective ozone layer. This realization – along with the shocking discovery in 1985 of a hole in the ozone layer over Antarctica – led nations to adopt the Montreal Protocol.

NOAA and AGAGE researchers meet regularly to discuss their findings, which they summarize in reports for the parties to the Montreal Protocol. These reports document the decline in the concentrations of CFCs in the atmosphere and they have identified other





**The Jungfrau research station in Switzerland is part of a global network that monitors the atmosphere.**

ozone-damaging chemicals. As such, scientists have continued to provide input into the protocol, which nations have updated to limit the production of other harmful gases. “It wasn’t a one-stop scientific treaty,” says David Fahey, an atmospheric chemist with NOAA, and one of the four co-chairs of the Scientific Assessment Panel of the Montreal Protocol.

The teams monitoring the air are forever playing catch up as new compounds appear in the skies. Even before CFCs were banned, manufacturers developed substitute coolants such as hydrochlorofluorocarbons (HCFCs). But researchers quickly found that these compounds also damage the ozone layer, and a 2007 amendment to the protocol called for the complete ban of production and consumption of HCFCs by 2030. Next came a third generation of coolant, the hydrofluorocarbons, or HFCs. These don’t contain chlorine or bromine, and so they don’t damage the ozone layer. But they turned out to be powerful greenhouse gases; most have a warming power between 1,400 and 5,000 times greater than that of carbon dioxide.

Consequently, in 2016, delegates agreed on the Kigali Amendment to the Montreal Protocol, which calls for cutting the production and use of HFCs by 80–85% by the late 2040s. The amendment entered into force at the start of 2019 with the goal of avoiding warming by up to 0.5 °C.

Monitoring stations such as Jungfrauoch track progress towards those goals in different parts of the world; sometimes they find problems. Scientists at the station found that northern Italy had emitted between 26 and 56 tonnes of HFC-23 per year in 2008–10, yet the official Italian inventory had estimated only 2.6 tonnes for the whole country.

### **Blindsided**

Until a few years ago, it seemed that the main threats to the ozone layer were on their way out and scientists could focus on the newer gases. Then came the first hints of trouble.

One day in 2013, Montzka ran the air from his weekly delivery of flasks through the mass spectrometer he had designed nearly 30 years earlier. But when he looked at the output of these routine measurements from the previous few months, he noticed something odd: the levels of CFC-11 were not declining as fast as before.

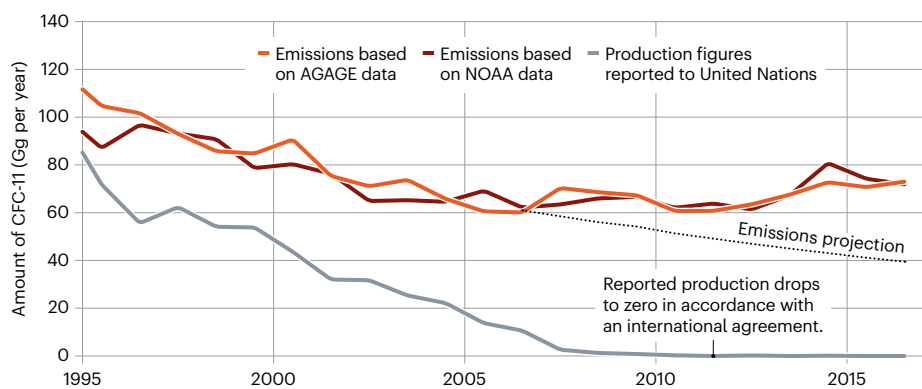
To Montzka, the observation made no sense — production of CFCs had been phased out worldwide three years earlier. Before 2012, the concentration of CFC-11 had dropped by about 0.8% per year, but Montzka’s flask data suggested the decline rate had slowed substantially. “I was totally amazed, I couldn’t believe it,” Montzka says. “Then I thought to myself

UDO BERNHART/DUMONT BILDARCHIV/PICTURE ALLIANCE



## SECRET STOCKS

Researchers use data from two air-monitoring networks to calculate emissions of CFC-11, which can come from new production or leakage from older products. Emissions declined as expected until 2005, but then plateaued and started to rise because of rogue manufacturing.



that it was just some blip that will go away next year – something weird has happened in the atmosphere, or in my instrument.”

Montzka double-checked his measurements and then, for the next few years, he and the international team searched for possible explanations. Eventually, the trail of evidence led to a single conclusion: emissions of CFC-11 were going up rather than down, pointing to a violation of the Montreal Protocol (see ‘Secret stocks’). “It did take a while to unravel the story in a way that I thought would be useful to the international community,” Montzka says.

Between 2002 and 2012, CFC-11 emissions averaged 54,000 tonnes per year, owing to gradual leakage of old stores of the compound contained in foam insulation and appliances made before the mid 1990s. But the researchers found that between 2014 and 2016, average emissions grew to 67,000 tonnes a year – an increase of roughly 25%<sup>1</sup>. They also noted that, in 2013, the flask data at the Mauna Loa Observatory in Hawaii suddenly showed increased levels of CFC-11 in the pollution plumes regularly recorded at the site. On closer investigation, they found that the sources of those plumes, and the uptick in CFC-11 emissions, came from eastern Asia.

A team of scientists immediately began to look for clues in an independent set of measurements, in particular those from the AGAGE stations on Jeju Island in South Korea, and Hateruma in Japan. Data from these stations revealed spikes in CFC-11 whenever plumes of pollution passed by. And the spikes had grown since 2013.

With this information, the scientists ran computer models using atmospheric circulation data and the monitoring-station measurements to determine where the pollution was coming from. Four independent modelling groups worked on solving the puzzle, and all came back with the same answer: about 7,000 tonnes per year were coming from the Chinese provinces of Shandong and Hebei<sup>2</sup>.

The newly discovered emissions will not significantly delay recovery of the ozone layer,

says Matthew Rigby, an atmospheric chemist at the University of Bristol, UK. “But if they carry on, we could be seeing delays of years or more,” he says.

## A close call

On 4 November 2019, Tina Birmipili, executive secretary of the UN Ozone Secretariat, delivered her opening speech at the 31st Meeting of the Parties to the Montreal Protocol in Rome. She began by praising the success of the treaty so far and the decisive action taken by China to address its emissions of CFC-11, including setting up a national monitoring network and increased penalties for companies that violate production bans. “CFC-11 was an alarm for all parties to ensure that they address illegal production swiftly and send a clear message to those who would break the law,” Birmipili says.

Then Birmipili turned her attention to some unanswered questions around the unexpected CFC-11 emissions. The researchers’ most recent published findings estimate that CFC-11 emissions from China account for 40–60% of the global increase between 2014 and 2017, but that leaves 4,000–10,000 tonnes unaccounted for<sup>2</sup>.

Right now, the researchers aren’t in a position to say whether there are other sources of illegal emissions or whether uncertainties in their models can account for the remaining percentage of the global trend, Rigby says. In the future, they will try to improve their models to see if they can glean a more accurate picture of the CFC-11 changes, he says. Montzka thinks that this time the monitoring community was lucky: researchers were able to detect the global trend change fairly early and happened to be making measurements near the region where at least some of the new emissions were coming from. But if CFC-11 had emanated from India, Russia or South America, the existing networks wouldn’t have been able to identify the location of the source because no regional stations exist nearby.

When Montzka stepped up to the podium in Rome, he presented some fresh observations from the global monitoring data. In 2018, the

rogue emissions seemed to slow or disappear. The decline of the global concentrations of CFC-11 accelerated, and the amount of the gas in plumes reaching the monitoring stations in Hawaii and Jeju Island substantially decreased. Although researchers have yet to fully check the latest measurements, they take heart from the trend. “The evidence suggests that the Montreal Protocol is being effective in yet another set of circumstances – in this case, unprecedented circumstances,” Fahey says.

If the CFC-11 concentrations continue to decline over the next few years, it will mark a significant victory for the scientists and their monitoring networks. “There’s always the discussion of whether it is really important that we are still here,” says Stefan Reimann, an atmospheric chemist at EMPA. “And, yes, history proves that we still have to be here.”

The rogue-emissions incident highlights weaknesses in the current system, which was developed to investigate the science of how the atmosphere is changing, not to track emissions, says geochemist Ray Weiss at the University of California, San Diego. “We never expected to see a violation, which is a lesson in itself really.”

In response to the latest challenge, NOAA added a flask-collection site on the west coast of South Korea to gather more information from eastern Asia. And this year, the parties will continue to discuss what is needed to ensure a similar violation doesn’t happen again, Birmipili says.

Meanwhile, the scientists are maintaining their strategy of watching, waiting and investigating. At Jungfraujoch, Vollmer is paying close attention to the latest generation of coolants: hydrofluoroolefins (HFOs). As those break down, some of them, such as one known as HFO-1234yf, can decompose into trifluoroacetic acid, which is toxic to some plants and soil organisms. The German and Norwegian environment agencies have recommended more research on the HFOs.

Measurements at Jungfraujoch show a rapid rise in these compounds. In 2011, HFO-1234yf appeared in none of Vollmer’s samples. By 2018, it was in 71% of them.

Currently, industry produces only a small amount of HFOs because the phase-out of HFCs has just begun. “But if you make a back-of-the-envelope calculation and you replace all the compounds that we’ve been using previously by the HFOs, there are going to be huge quantities of these gases,” Vollmer says.

So he makes the journey each month to the high, glaciated saddle between two peaks in the Alps, where Jungfraujoch’s instruments hum away day and night. “We have to keep watching,” he says.

**Jane Palmer** is a freelance writer based in Colorado.

1. Montzka, S. A. *et al. Nature* **557**, 413–417 (2018).
2. Rigby, M. *et al. Nature* **569**, 546–550 (2019).



# Books & arts



UNDERWOOD ARCHIVES/UIG/SHUTTERSTOCK

US Army nurses in 1947. Shifting social norms have driven the swift rise and demise of smoking in many places.

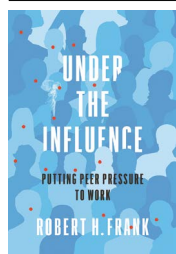
## Peer pressure shapes our world

Social context affects our actions. Policymakers should leverage that to cut emissions, boost health and more, a book argues. **By Thomas Dietz**

In 1989, just 12% of US adults favoured legalization of same-sex marriage; by 2015, that figure was around 60%. What triggered the transformation? In *Under the Influence*, economist Robert Frank reveals that peer pressure lies behind many such step changes. Once views began to shift, the process was self-reinforcing.

As Frank drives home, we humans are especially adept at learning from our peers. Our decisions are strongly influenced by social norms – what we think others are doing, and what we think they think we

should do. In some circumstances, we can be self-interested; in others, we can be altruistic. So it's not surprising that much



**Under the Influence:  
Putting Peer Pressure  
to Work**

Robert H. Frank  
Princeton Univ. Press  
(2020)

of social-science research focuses on social context in decision-making. Frank reviews extensive evidence from studies across a number of disciplines on how peer pressure shapes the dynamics of smoking, drinking, obesity, consumerism and many other important social issues.

### Pressure point

Because the tendency to emulate can lead to rapid social change, for better or worse, it is a key lever for policy. Yet, asserts Frank, that message has yet to reach many policy analysts

and economists. *Under the Influence* offers a corrective through compelling arguments for incorporating social contexts into the design of policy on climate change, public health, the financing of public goods, social justice, taxation and beyond.

Among the cascades of change Frank examines are ‘arms races’, which can focus on anything from nuclear weapons to consumer goods. They are a type of commons dilemma or collective-action problem: the pursuit of narrow self-interest leads to overuse of a resource, and disaster. (If foresters, for instance, limit the number of trees they fell every year, the forest can regenerate, to the benefit of all; if they each boost their own short-term profits by maximizing their felling, the forest ecosystem might collapse.) But in an arms race, what matters is not your absolute measure of resources. It is what you have compared with what I have. Thus, everyone has an incentive to accumulate resources in a never-ending upward spiral.

### Boom and bust

Frank points, for example, to the sharp increases in US housing prices that led to the bubble of the early 2000s. To ensure access to the best school districts, buyers competed to live in the most affluent neighbourhoods, bidding up housing costs inexorably. The result was unrealistic prices, unsustainable mortgage burdens and a slump in price that led to bankruptcies and the collapse of lenders – all of which contributed to the 2008 economic meltdown.

Frank examines another problematic arms race: the widespread opposition of the rich to increased taxation. This, he argues, hinges on what he calls the “mother of all cognitive illusions”: the belief that happiness is based on absolute wealth (and spending power), which higher taxes would slash. Frank counters that view, asserting that rich people’s well-being is based on relative wealth – their position compared with that of their peers. A tax affecting all top earners would maintain relative position, whatever the effect on absolute spending power. His analysis is timely,

**“In an arms race, what matters is not your absolute resources. It is what you have compared with what I have.”**

because low and declining US tax rates for the top income bracket have led to a loss of government revenue and, in turn, massive underinvestment in public goods such as education and infrastructure. Frank suggests a remedy: taxing consumption (income minus savings) for the wealthiest.

One of the great strengths of *Under the Influence* is Frank’s use of research from across the social sciences, including psychology and political science. Yet he fails to engage with much that’s salient to his arguments here. For instance, regarding policy challenges such as climate change and obesity, he admits that his “deepest passion” is efficiency – that is,

he favours taxation over regulation. Thus, he adopts a standard utilitarian approach to decision-making. To demonstrate the success of this approach, he cites the US policy that placed a price on sulfur dioxide emissions from 1995, significantly reducing levels of acid rain. But when he discusses the importance of in-depth deliberation in resolving conflicts, and in changing individual views on gay rights and environmental protection, he does not mention the extensive literature on how deliberative processes can underpin good decision-making, a theory complementary to his utilitarianism.

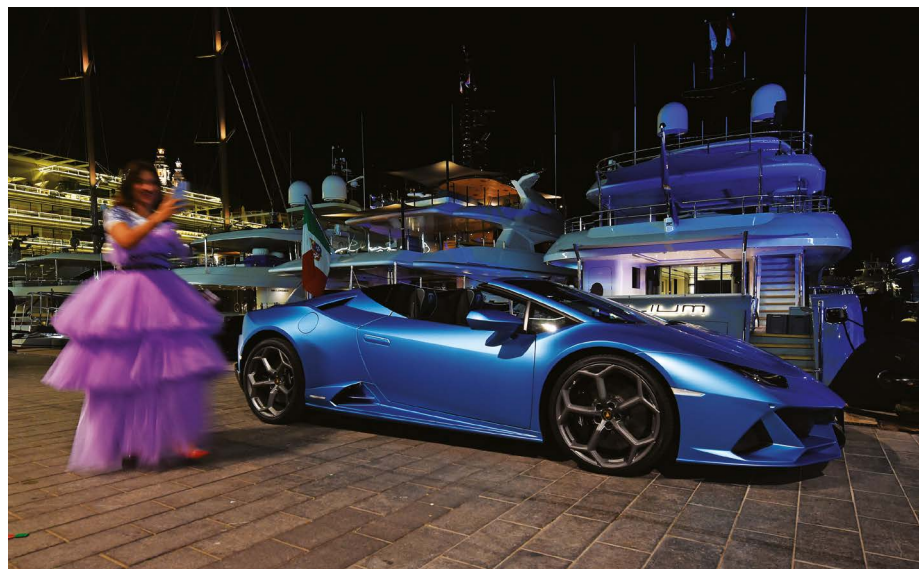
### Unexplored factors

Frank’s analysis would thus benefit from even deeper digging into findings on context, social structure, power and social interaction, such as the critique of growth dynamics in environmental sociology or the 2017 book *Beyond Politics*, an analysis of private environmental governance by Michael Vandenbergh and Jonathan Gilligan. For example, Frank’s argument about the well-being of the affluent resting on relative status does not factor in the possibility that rich people might be seeking political power and influence on government instead. Among the richest, power might depend on absolute wealth. Similarly, his thoughtful chapter on climate change does not fully address opposition to climate policy from powerful fossil-fuel interests.

Moreover, Frank mentions only in passing issues such as the human tendency to associate with those like us (homophily) and to affirm what we already believe (confirmation bias). In the social networks of government officials, lobbyists and others who influence policy, these tendencies lead to polarization and a lack of action on serious problems. So although Frank urges us to consider context, he misses the need to pay more attention to the structure of contexts, including inequality and power.

Of course, one book, however broad its compass, cannot cover everything. And even where I felt Frank had not tackled important lines of research, those gaps point to the need to think more deeply about human actions and the policies that shape them. At a time of multiple impending crises, *Under the Influence* will provoke your thinking in constructive ways.

**Thomas Dietz** is university distinguished professor in sociology and of environmental science and policy at Michigan State University in East Lansing.  
e-mail: tdietzvt@gmail.com



For many people, wealth relative to others is more important than absolute spending power.



# Comment



FRANCOIS MORI/AP/SHUTTERSTOCK

Christiana Figueres at the 21st United Nations Climate Change Conference. She led the negotiations that produced the 2015 Paris agreement.

## The secret to tackling climate change

Christiana Figueres

To the world leaders mustering in Davos: set your minds to reaching net-zero emissions, and you can forge the future we need.

**A**s political leaders, industry executives and celebrities gather this week for their yearly networking meeting in Davos, Switzerland, top of their agenda is the need to halve global carbon emissions by 2030.

Of the many barriers to achieving this goal, the greatest is mindset. I had to learn this a decade ago when I was appointed to lead the international climate-change negotiations that resulted in the 2015 Paris agreement: ultimately, 195 nations pledged to reduce emissions and alter their economies to protect our planet. They also agreed to increase

their efforts towards net-zero emissions substantially every five years. That makes 2020 a crucial year. We cannot afford for governments to let that key commitment slip.

The Paris agreement was a breakthrough after a devastating collapse in Copenhagen in 2009, when years of preparation and two weeks of excruciating around-the-clock negotiations produced only a weak, legally irrelevant accord. Copenhagen was a free-for-all of political frustration, outrage and disagreement – with the global north and global south set against each other. Last month's United Nations climate meeting in Madrid left many of us similarly bereft. That makes the lesson of how we got from Copenhagen to Paris all the more relevant.

It started with my making a big mistake in the summer of 2010, at a press conference with 40 journalists in a windowless room at the Maritim Hotel in Bonn, Germany. When asked whether a global agreement on climate change would ever be possible, I blurted out, without thinking, what most already thought: "Not in my lifetime." That's how close I came to

making failure a self-fulfilling prophecy.

I immediately realized that, before we could consider the political, technical and legal parameters of an eventual agreement, I had to dedicate myself to changing the mood: there could be no victory without optimism. I decided to set a clear intention: even if we did not know precisely how, a global deal would emerge, simply because it was necessary. It was that contagious frame of mind that led to effective decision-making, despite the enormous complexities under which we were operating. When the Paris agreement was achieved, the optimism that people felt about the future was palpable – but, in fact, optimism had been the primary input.

Since then, science has become clearer about the threats of climate change: now, even our children know that business as usual will lead to destroyed infrastructure, devastating loss of plants and animals, and millions of people struggling in regions made uninhabitable from rising temperatures and lack of fresh water.

What is much less clear is what life will look like in those places where we do what is necessary to limit warming to 1.5 °C, as stipulated by the Paris agreement. To get to what we achieved in Paris, we moved away from confrontational blaming-and-shaming to appreciating shared opportunities. Now, we must picture, say, cities full of green spaces pulling carbon dioxide from the atmosphere; widespread public transport; thriving wildernesses; rural economies rebooted for sustainable agriculture; and jobs in renewable-energy projects.

Optimism is about acknowledging difficulties – and losses – yet still designing a better future. An excellent example is the European Union's proposed European Green Deal, announced in December 2019. This explicitly reframes an urgent challenge as a unique opportunity to create a “resource-efficient and competitive economy” that will generate jobs, purify air and mobilize industry, agriculture and other sectors to deliver net-zero emissions by 2050.

My own country, Costa Rica, has already launched an economy-wide plan to ‘decarbonize’ by 2050. This ambitious plan, the first of its kind when it was announced last February, will expand forests and promote electric taxis and public buses. It is based on respect for human rights and gender equity, and clearly recognizes the opportunity for decarbonization to revitalize the economy.

Most executives already understand that they need to contribute to climate stabilization just to ensure that their businesses have a future. The number of companies



Costa Rica has launched a decarbonization plan that will expand the country's forest cover.

setting science-based targets in line with a 1.5 °C trajectory doubled between September and December last year. Similarly, the combined assets managed by the Net-Zero Asset Owner Alliance – a group of investors aligning their portfolios with a 1.5 °C future – had surged from US\$2.4 trillion to \$4 trillion within two months of its launch in September 2019. Leaders in the oil and gas industries have told me privately that shareholder and public pressure, plus questions from their own children, have prompted them to shift their practices.

Despite this, I posit that most people, including many of those attending the Davos meeting, still harbour the view that it is impossible to truly transform our economy in one decade. We cannot afford such fatalism. Swift change has happened before, and without being driven by planetary necessity: the global Internet is just 30 years old.

If we can see where we are going – a future in which humanity does what is necessary to preserve the planet as we know and love it – we will take faster, surer steps to get there. That visualization is all the more important because how we are going to get to this future will feel unfamiliar. The transition of technologies and systems in music and information makes sense only because we have seen vinyl records yield to streaming services and paper superseded by mobile multimedia. We must be ready to shape the necessary transition for energy, transport and more. And we must understand that this transition will be driven collectively.

The global economy is a huge, complex system. As I learnt during my stewardship of

the Paris agreement, if you do not control the complex landscape of a challenge (and you rarely do), the most powerful thing you can do is to change how you behave in that landscape, using yourself as a catalyst for overall change.

Imagine a person who wants to run a marathon and then concentrates on the fact that they can't yet even run a mile: they begin to close the space of possibility. But, if that person adopts a different mindset, commits to a training schedule and visualizes passing the finish line, their goal is much more likely to be achieved.

To all the people gathering in Davos, and all those watching from the outside, I urge you to move firmly into a state of stubborn optimism. The Anthropocene, the proposed geological age we now live in, does not need to go down in history as the age characterized by human-induced destruction. It can be the time when we rewrite our expected future for a better one: we still hold the pen. We must conceive of success and take immediate steps towards it.

## The author

**Christiana Figueres** was the executive secretary of the United Nations Framework Convention on Climate Change from 2010 to 2016. She is a co-founder of Global Optimism, an enterprise that aims to stimulate social and environmental change, and co-author of the forthcoming book *The Future We Choose*. e-mail: cfigueres@mission2020.global



# Correspondence

## Curb spread of virus emerging in China

I applaud Chinese colleagues' prompt release of the genome sequence of the virus responsible for the mystery respiratory illness in Wuhan in central China (see *Nature* <http://doi.org/djhc>; 2020). The agent is a previously unknown type of coronavirus that is distantly related to the severe acute respiratory syndrome (SARS) coronavirus. To curb the spread of the virus, its animal reservoir must be quickly identified and human-to-human transmission thoroughly investigated (see also [go.nature.com/2ua489i](http://go.nature.com/2ua489i)).

The authorities have been understandably cautious after the early misidentification of the SARS pathogen in 2003. However, the results of animal testing from a seafood market in Wuhan, where the virus was initially isolated, must be released as soon as possible. The virology community also feels that human-to-human transmission should not be ruled out without compelling evidence.

This information is particularly crucial because tens of millions of people will be travelling – and consuming potentially contaminated animal meat – to celebrate the Chinese New Year on 25 January. The public needs clear instructions and guidance.

Controlling the spread of emerging and re-emerging viruses calls for international efforts. China's research collaborations and data-sharing must continue – including with the United States, despite other problems with their relations.

**Shan-Lu Liu** The Ohio State University, Columbus, Ohio, USA.  
[liu.6244@osu.edu](mailto:liu.6244@osu.edu)

## Grants: don't leave it to luck

I was shocked to read that a growing number of funding bodies are assigning research grants randomly (*Nature* 575, 574–575; 2019). As an early-career researcher, I might be expected to gain from such a system, given that I could land a windfall without having my case judged against the competition. But I want my career to be built on achievement, as recognized and promoted through conventional grant awards – not undermined by a lottery system.

Some researchers might see random funding as more time-efficient, because it dispenses with the review process. It spares reviewers the burden of differentiating between the lowest-ranked successful candidate and the highest-ranked candidate who didn't make the cut. However, for a researcher just starting out, a positive review based on the applicant's contributions to the literature and other scientific merits is crucial for advancement.

And if lottery-based grants become widespread, academic research will suffer as fruitful ideas are arbitrarily stalled. Leaving success up to lady luck is not a solution.

**Howard Vindin** University of Sydney, Australia.  
[hvin6646@uni.sydney.edu.au](mailto:hvin6646@uni.sydney.edu.au)

## Grants: lottery is laziness

The idea of a funding lottery (*Nature* 575, 574–575; 2019) is, in my view, a classic bureaucratic response to a process that bureaucracy finds too hard to handle.

The review of scientific grant applications depends on an assessment of their quality, requiring a strict combination of evidence and intellectual judgement. Stuff that, say the bureaucrats. "Let's make it a lottery, and save ourselves time and money." Sure, some applications might flourish that otherwise would not, but what about the high-quality research that has been carefully constructed over time and is suddenly de-funded? Such a funding system is, in effect, anti-intellectual. It is a research version of publication bibliometrics that focus merely on citation counts, not on quality.

Academia must resist this bureaucratization of research and publishing by well-meaning but scientifically inept bureaucrats. Otherwise, science itself stands to be plunged into the same miasma of metrics and bureaucracy-benefiting processes that have already weakened other great institutions, many examples of which are described in Jerry Muller's book *The Tyranny of Metrics* (see *Nature* 554, 167; 2018).

**Andrew Beattie** Macquarie University, Sydney, Australia.  
[andrew.beattie@mq.edu.au](mailto:andrew.beattie@mq.edu.au)

## Climate change: be mindful at meetings

Scientists are keen to lower the toll their work takes on the planet (see, for example, O. Hamant *et al.* *Nature* 573, 451–452; 2019). At a recent Harvard conference on sociology and climate change, Hannah Holleman – a sociologist at Amherst College in Massachusetts – offered us a gentle reminder of how our research is embedded in everyday practices (see [go.nature.com/3acmulr](http://go.nature.com/3acmulr)).

In her memorable opening statement, Holleman drew attention to the debt we owe to the native peoples whose traditional homelands are now occupied by the university, the natural resources used to build the venue, the production of sustenance for the event, and the fossil fuel needed for us to convene. She pointed out that the organic materials used would return, as waste, to the land.

This unusual opening to an academic discussion landed a strong emotional punch. It was a powerful reminder – even for scholars who are well informed and deeply committed to solving the biodiversity and climate crises – of our shared responsibility and accountability. It used mindfulness as a way to amplify the urgency of that message. This approach could bear further exploration at other meetings on climate change.

**Charles Davis** Harvard University Herbaria, Cambridge, Massachusetts, USA.  
[cdavis@oeb.harvard.edu](mailto:cdavis@oeb.harvard.edu)

**Benjamin Goulet-Scott** Arnold Arboretum, Boston, Massachusetts, USA.

**Jason Beckfield** Harvard University, Cambridge, Massachusetts, USA.

# News & views



Figure 1 | The Ganges river delta.

## Hydrology

# The changing shapes of river deltas

Nick van de Giesen

A model has been devised that quantitatively describes how the shape of a river delta is affected by sediments, tides and waves. It reveals that the area of delta land is increasing globally, as a result of human activities upstream. **See p.514**

Undisturbed river deltas are diverse ecosystems that encompass tidal wetlands and floodplains. Because of their rich soils and convenient positions for trade and transport, many deltas have also become hotspots of socio-economic development. The Nile delta, for example, with its iconic triangular

shape, has been one such locus for more than 5,000 years. Not all deltas are triangular, however – their morphology can vary widely. On page 514, Nienhuis *et al.*<sup>1</sup> report a model that correlates the forces that shape deltas with delta morphology, and use it to analyse the shapes of some 11,000 coastal deltas. This

global overview allows the authors to assess how delta morphology is affected by changes in sediment delivery caused by river damming and soil erosion.

The authors' model estimates delta morphology on the basis of a quantitative characterization of three main drivers that shape deltas. These are: sediment delivered by the river; wave action that redistributes sediment along the coast; and sediment transported into or out of the delta by tidal flows. The relative influences of these drivers were used to determine two key morphological metrics; namely, the protrusion of the delta into the sea and the shape of the river channel. For example, Nienhuis *et al.* infer from the model that when the effects of sediment delivered by the river are greater than the effects of wave action, deltas protrude relatively far into the sea. Alternatively, the authors conclude that deltas widen towards the sea into a trumpet shape when tidal flows are important and sediment delivery is low. Nienhuis *et al.* validated their model by comparing the projected morphologies with those of real deltas, and



provide robust statistics on the reliability of the results, which is a key strength of the study.

Note that the authors' definition of what constitutes a delta is broad (see the Methods section of the paper for the criteria used), which means that their model is truly global. However, the model's ability to capture the general behaviour of all deltas comes at the expense of fine-grained accuracy – there will almost inevitably be errors in the morphologies projected for some individual deltas. Nevertheless, the model's results are statistically valid at a global level.

Nienhuis and colleagues used their model to estimate the effects of upstream human interventions on delta morphology during the period 1985–2015. They found that dam building led to decreases in sediment delivery, whereas accelerated soil erosion caused by deforestation increased sediment delivery. Of the approximately 11,000 deltas analysed, about 9% are significantly affected by reduced sediment delivery, producing a total land loss of 127 square kilometres per year, whereas about 14% received increased sediment, causing a total gain of 181 km<sup>2</sup> yr<sup>-1</sup> during the study period. The reason more deltas have experienced an increase in sediment delivery, rather than a decrease, is simply that the effects of massive deforestation have outpaced sediment trapping by dams.

Previously reported state-of-the-art studies<sup>2,3</sup> of global coastal morphology involved the computationally intensive analysis of extremely large archives of satellite images, which have become available in the past few years. These studies also revealed a net increase in land surface area. Many of the land gains could be explained by large-scale phenomena, such as the disappearance of the Aral Sea in central Asia, and by extensive land-reclamation projects along the China coast. But beyond those special cases, it is also crucial to learn in greater detail where and why river deltas have gained or lost land across the globe. Nienhuis *et al.* fill in this key part of the puzzle.

The new study also reveals notable regional patterns. For example, arctic river deltas have seen almost no change in morphology. Sediment delivery by rivers in North America has fallen overall, leading to large land losses – in the Mississippi delta, for example. And the largest land gains are in eastern South America and in south, southeast and east Asia, where soil erosion due to deforestation has caused a net growth in delta areas, despite the construction of sizeable dams in these regions.

Large deltas, such as those of the Niger, Huang He and Mekong, have great socio-economic value. Such densely inhabited deltas typically experience many pressures in addition to changes in sediment delivery, such as stresses associated with groundwater pumping, sand mining, dyke construction and loss

of biodiversity<sup>4–6</sup>. For these highly complex deltaic systems, local studies will be needed to assess the problems that adversely affect their morphology and to define specific solutions<sup>6</sup>. However, most of the deltas considered by Nienhuis and co-workers are much smaller. This could skew the picture painted by the overall numerical results, because large deltas have a much greater global impact than do small ones, but represent a tiny fraction of the total number of deltas analysed in the study. For example, the study calculates that the net land gain for all deltas was 54 km<sup>2</sup> yr<sup>-1</sup> during the period studied, which seems like good news. But this area is tiny compared with the 105,000 km<sup>2</sup> covered by the Ganges delta alone (Fig. 1) – which, with its population of 170 million people, is subject to a multitude of stresses<sup>7</sup>. We should therefore not be complacent about the new findings.

Nienhuis *et al.* did not include sea-level rise in their model, but sea levels rose by about 10 cm over the period studied (see [go.nature.com/2tpjpxg](https://go.nature.com/2tpjpxg)). This will probably not have produced observable losses of delta land, given the large spatial variability of sea-level rises. Nevertheless, it would be interesting to see whether measurable losses did occur. The authors' model provides a useful description of the background dynamics of changes in delta morphology against which the impact of rising seas can be measured once sea levels approach predicted increases of 60 cm (ref. 8) or more<sup>9</sup>, as a result of global warming. Severe sea-level rise will undoubtedly cause coastline recession in deltas, as it has in the geological past<sup>10</sup>.

Validated global models describing key parts of the Earth system are crucial in this time of unprecedented human-induced climate change. Deltas connect the terrestrial and maritime branches of the hydrological cycle and the associated sediment fluxes. As such, they encapsulate many key indicators of global change. By accounting for the baseline effects on deltas of human activities such as dam building and deforestation, Nienhuis and colleagues have provided a fundamental framework that will help assessments of the impacts of climate change for decades to come.

**Nick van de Giesen** is in the Department of Water Management, Delft University of Technology, 2628 Delft, the Netherlands. e-mail: [n.c.vandegiesen@tudelft.nl](mailto:n.c.vandegiesen@tudelft.nl)

1. Nienhuis, J. H. *et al.* *Nature* **577**, 514–518 (2020).
2. Pekel, J.-F., Cottam, A., Gorelick, N. & Belward, A. S. *Nature* **540**, 418–422 (2016).
3. Donchyts, G. *et al.* *Nature Clim. Change* **6**, 810–813 (2016).
4. Renaud, F. G. *et al.* *Curr. Opin. Environ. Sustain.* **5**, 644–654 (2013).
5. Tessler, Z. D. *et al.* *Science* **349**, 638–643 (2015).
6. Bucx, T., Marchand, M., Makaske, B. & van de Guchte, C. *Comparative Assessment of the Vulnerability and Resilience of 10 Deltas – Synthesis Report*. Delta Alliance Rep. 1 (2010); [go.nature.com/2ssuqhx](https://go.nature.com/2ssuqhx)
7. Auerbach, L. W. *et al.* *Nature Clim. Change* **5**, 153–157 (2015).
8. Intergovernmental Panel on Climate Change. *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (eds Core Writing Team, Pachauri, R. K. & Meyer, L. A.) p.60 (Table 2.1) (IPCC, 2014).
9. Garner, A. J. *et al.* *Earth's Future* **6**, 1603–1615 (2018).
10. Smith, D. E., Harrison, S., Firth, C. R. & Jordan, J. T. *Quat. Sci. Rev.* **30**, 1846–1860 (2011).

## Cancer immunology

# B cells to the forefront of immunotherapy

**Tullia C. Bruno**

Three studies reveal that the presence in tumours of two key immune components – B cells and tertiary lymphoid structures – is associated with favourable outcomes when individuals undergo immunotherapy. **See p.549, p.556 & p.561**

Current immunotherapies aim to reinvigorate immune cells called killer T cells to fight cancer, but only 20% of individuals with the disease see a lasting clinical benefit from this type of treatment<sup>1</sup>. Focusing on other immune cells in patients' tumours might help us to improve these outcomes. Three studies, by Cabrita *et al.*<sup>2</sup> (page 561), Petitprez *et al.*<sup>3</sup> (page 556) and Helmink *et al.*<sup>4</sup> (page 549), now demonstrate that the presence of B cells

in human tumours in compartments called tertiary lymphoid structures (TLS) is associated with a favourable response to immunotherapy. These complementary studies add to the immunotherapy toolbox by providing new ways of predicting prognosis.

The presence of B cells in tumours has been considered to be a predictor of increased patient survival<sup>5,6</sup>, but there are reports of both anti- and pro-tumour roles for B cells<sup>7</sup>.

# News & views

## Cancer immunology

# B cells to the forefront of immunotherapy

Tullia C. Bruno

Three studies reveal that the presence in tumours of two key immune components – B cells and tertiary lymphoid structures – is associated with favourable outcomes when individuals undergo immunotherapy.

Current immunotherapies aim to reinvigorate immune cells called killer T cells to fight cancer, but only 20% of individuals with the disease see a lasting clinical benefit from this type of treatment<sup>1</sup>. Focusing on other immune cells in patients' tumours might help us to improve these outcomes. Three studies in *Nature*, by Cabrita *et al.*<sup>2</sup>, Petitprez *et al.*<sup>3</sup> and Helmink *et al.*<sup>4</sup>, now demonstrate that the presence of B cells in human tumours in compartments called tertiary lymphoid structures (TLS) is associated with a favourable response to immunotherapy. These complementary studies add to the immunotherapy toolbox by providing new ways of predicting prognosis.

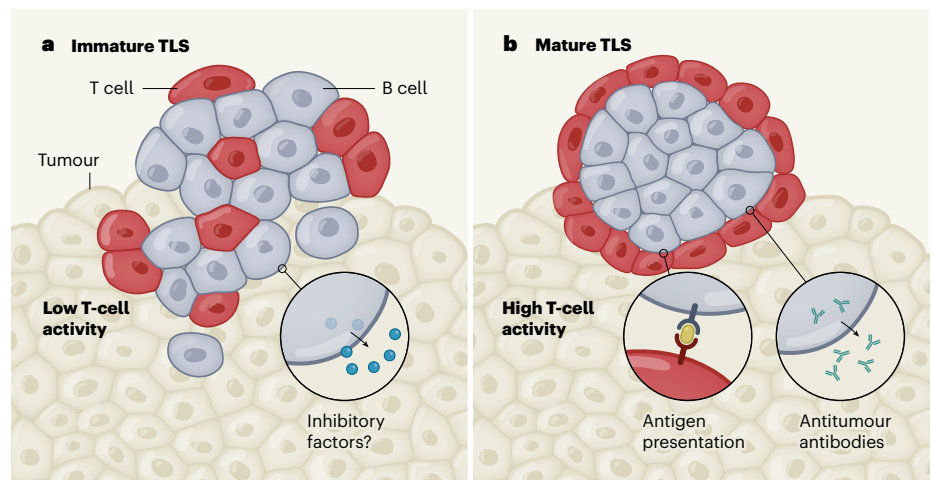
The presence of B cells in tumours has been considered to be a predictor of increased patient survival<sup>5,6</sup>, but there are reports of both anti- and pro-tumour roles for B cells<sup>7</sup>. These differing reports reflect the multiple roles that B cells can have in tumours. One component of the antitumour function of B cells is B-cell activation. This process involves the binding of tumour-derived proteins to the B-cell receptor protein on the cell surface and the subsequent processing of these tumour-derived proteins into smaller fragments called antigens. Further co-factors are also involved in activation. Activated B cells can release antibodies that tag tumour cells for attack by other cellular players of the immune system (a process known as antibody-dependent cell death)<sup>8</sup>, and can 'educate' T cells by presenting them with tumour antigens, enabling the T cells to target tumour cells effectively<sup>9</sup>. However, B cells in tumours can produce inhibitory factors that hinder the function of immune cells (Fig. 1). These might be signalling molecules that suppress the immune system<sup>7,10,11</sup> or

inhibitory molecules on the surfaces of B cells that limit the body's ability to target and kill tumour cells.

TLS are aggregates of immune cells (mostly T and B cells) that arise in response to immunological stimuli. Mature TLS nurture B-cell development and function in an inner region of the structure called the germinal centre, whereas immature TLS do not contain proper germinal centres, and might not nurture full B-cell function. The presence of TLS

in a tumour also correlates with increased patient survival in many cancer types<sup>12</sup>. The three current studies confirm this trend in the context of immunotherapy, demonstrating that infiltration of B cells into a tumour, along with the presence of TLS, is associated with an improved response to this type of treatment.

Cabrita *et al.* studied individuals who had a type of cancer called metastatic melanoma, and Petitprez *et al.* investigated people with sarcoma, a cancer of the bone. Both teams found that the presence of B cells in TLS in the tumour before treatment was associated with an increased chance that patients' tumours would respond to immunotherapy. Helmink *et al.* corroborated these findings for metastatic melanoma, and reported the same pretreatment trend in renal cell carcinoma. These authors also demonstrated that, during treatment, TLS are more prevalent in people who have tumours that are responding to treatment than in those whose tumours are not. This timing is important – when present before treatment, TLS could be considered a predictor of patient response to immunotherapy, whereas the presence of TLS during treatment indicates that key combinations of immune cells are being manipulated to induce TLS formation. Identifying these cell



**Figure 1 | Multifaceted B cells in the tumour microenvironment.** B cells are thought to have multiple roles in suppressing or promoting the immune system's ability to kill tumour cells, depending on whether they are located in immature or mature compartments called tertiary lymphoid structures (TLS), which also contain T cells. **a**, In poorly structured, immature TLS, one hypothesis is that B cells generate inhibitory factors. These might be molecules released from B cells that dampen the response of other immune cells, or molecules on the surfaces of B cells that hinder the targeting and destruction of tumour cells. Both of these inhibitory mechanisms might arise if B cells have less interaction with T cells and more interaction with the malignant tumour. Three studies<sup>2–4</sup> now provide indirect evidence that immature TLS are associated with low activity of T cells in tumours. **b**, By contrast, B cells in well-structured, mature TLS can release antibodies that could target tumours, and B cells can present a tumour-derived protein called an antigen (yellow) to T cells in the tumour, activating the T cells. The studies suggest that the presence of B cells in mature TLS is correlated with increased T-cell activity, improving the immune system's ability to target tumour cells, and increasing the likelihood that the tumour will respond to immunotherapy.



combinations could help in establishing new and effective immune-based therapies.

The three groups found that the B-cell and TLS signature was often more pronounced in responders than in non-responders. Furthermore, the signature was more prominent than typical T-cell signatures currently used for understanding immunotherapy outcomes. This suggests that B cells and TLS could have a key role in antitumour immunity.

In addition to these synergistic results, each study highlights a unique role for B cells or TLS in antitumour immunity. First, Cabrita *et al.* demonstrate that B cells in TLS synergize with killer T cells that could ultimately target tumour cells. Second, Petitprez *et al.* describe signatures characteristic of mature TLS in sarcoma. This implies that mature TLS can exist in tumour sites that are not normally thought to be infiltrated by immune cells, a phenomenon that has not previously been shown. Third, Helmink *et al.* find increased diversity of B-cell receptors in responders compared with non-responders. This indicates that pools of B cells in responders might have a greater ability to specifically recognize tumour antigens than do the B cells of non-responders.

These papers are technologically savvy, use patient populations that are statistically robust and bring B cells and TLS to the forefront of antitumour immunity. However, there is much still to learn. First, more emphasis should be placed on understanding how TLS form in tumours. It is clear that these structures are variable, and can be immature or mature. What does this diversity mean for the function of B cells in TLS, and what causes the induction of one 'flavour' of TLS versus another? The contribution of environmental factors such as smoking or viral and bacterial infections should be considered, along with a

person's gender, age and tumour type.

Researchers should also ask whether mature TLS could be routinely induced to form in tumours, to maximize B-cell immunity. Addressing this issue will require investigation of B cells and TLS in individuals who have not yet undergone treatment, as well as proper modelling of the human tumour microenvironment. Current evidence indicates that B cells actually impede antitumour responses in most mouse models of cancer<sup>13–15</sup>. However, TLS formation is rare in these animals, and a lack of TLS might alter the fate and subsequent function of B cells. Indeed, more knowledge about B-cell function outside TLS is needed to provide a complete picture of B cells in the tumour microenvironment.

There is still a need to define the full range of functions that B cells perform in tumours. In addition to their known roles in producing tumour-specific antibodies and presenting antigens<sup>8,9</sup>, B cells are likely to have other functions – for instance, inducing antibody-dependent cell death<sup>8</sup>. It will also be necessary to link these functions to specific B-cell types and to determine whether such cells are found inside or outside TLS. There are clear biomarkers for B-cell subsets, but linking these subsets to functions in human tumours would allow us to design treatments that optimize specific antitumour activities. Furthermore, this knowledge would help us to understand whether subsets of B cells perform separate tasks, or if there is crosstalk between subsets. For example, can the same B cell both produce a tumour-specific antibody and present antigens to T cells? Some of these studies can be done in human tumours, but in-depth mechanistic studies will require physiologically relevant models that contain naturally occurring TLS.

With regard to clinical implications, the

current studies suggest that therapeutics to enhance B-cell responses should be prioritized as a complement to T-cell-mediated immunotherapies. Researchers should now ask whether B cells could be engineered to target specific tumour antigens, similar to current efforts to engineer antigen-targeting T cells. More generally, could immunotherapies be improved by inducing B cells to form in TLS after a person has received T-cell-based immunotherapy?

Overall, the current studies should act as a springboard for future mechanistic studies of B cells and TLS in cancer. Understanding how current therapies can be combined with approaches to harness B cells and TLS will be crucial for the development of effective B-cell-specific immunotherapies.

**Tullia C. Bruno** is in the Department of Immunology, University of Pittsburgh, Pittsburgh, Pennsylvania 15215, USA, and at the UPMC Hillman Cancer Centre, Pittsburgh. e-mail: tbruno@pitt.edu

1. Brahmer, J. R. *et al.* *J. Clin. Oncol.* **28**, 3167–3175 (2010).
2. Cabrita, R. *et al.* *Nature* <https://doi.org/10.1038/s41586-019-1914-8> (2020).
3. Petitprez, F. *et al.* *Nature* <https://doi.org/10.1038/s41586-019-1906-8> (2020).
4. Helmink, B. A. *et al.* *Nature* <https://doi.org/10.1038/s41586-019-1922-8> (2020).
5. Shimabukuro-Vornhagen, A. *et al.* *Oncotarget* **5**, 4651–4664 (2014).
6. Germain, C. *et al.* *Am. J. Respir. Crit. Care Med.* **189**, 832–844 (2014).
7. Shalapour, S. *et al.* *Nature* **521**, 94–98 (2015).
8. DeFalco, J. *et al.* *Clin. Immunol.* **187**, 37–45 (2018).
9. Bruno, T. C. *et al.* *Cancer Immunol. Res.* **5**, 898–907 (2017).
10. Kessel, A. *et al.* *Autoimmun. Rev.* **11**, 670–677 (2012).
11. Khan, A. R. *et al.* *Nature Commun.* **6**, 5997 (2015).
12. Sautès-Fridman, C., Petitprez, F., Calderaro, J. & Fridman, W. H. *Nature Rev. Cancer* **19**, 307–325 (2019).
13. Affara, N. I. *et al.* *Cancer Cell* **25**, 809–821 (2014).
14. Shalapour, S. *et al.* *Nature* **551**, 340–345 (2017).
15. Ammirante, M. *et al.* *Nature* **464**, 302–305 (2010).

that optimize specific antitumour activities. Furthermore, this knowledge would help us to understand whether subsets of B cells perform separate tasks, or if there is crosstalk between subsets. For example, can the same B cell both produce a tumour-specific antibody and present antigens to T cells? Some of these studies can be done in human tumours, but in-depth mechanistic studies will require physiologically relevant models that contain naturally occurring TLS.

With regard to clinical implications, the current studies suggest that therapeutics to enhance B-cell responses should be prioritized as a complement to T-cell-mediated immunotherapies. Researchers should now ask whether B cells could be engineered to target specific tumour antigens, similar to

current efforts to engineer antigen-targeting T cells. More generally, could immunotherapies be improved by inducing B cells to form in TLS after a person has received T-cell-based immunotherapy?

Overall, the current studies should act as a springboard for future mechanistic studies of B cells and TLS in cancer. Understanding how current therapies can be combined with approaches to harness B cells and TLS will be crucial for the development of effective B-cell-specific immunotherapies.

**Tullia C. Bruno** is in the Department of Immunology, University of Pittsburgh, Pittsburgh, Pennsylvania 15215, USA, and at the UPMC Hillman Cancer Centre, Pittsburgh. e-mail: tbruno@pitt.edu

1. Brahmer, J. R. et al. *J. Clin. Oncol.* **28**, 3167–3175 (2010).
2. Cabrita, R. et al. *Nature* **577**, 561–565 (2020).
3. Petitprez, F. et al. *Nature* **577**, 556–560 (2020).
4. Helmink, B. A. et al. *Nature* **577**, 549–555 (2020).
5. Shimabukuro-Vornhagen, A. et al. *Oncotarget* **5**, 4651–4664 (2014).
6. Germain, C. et al. *Am. J. Respir. Crit. Care Med.* **189**, 832–844 (2014).
7. Shalapour, S. et al. *Nature* **521**, 94–98 (2015).
8. DeFalco, J. et al. *Clin. Immunol.* **187**, 37–45 (2018).
9. Bruno, T. C. et al. *Cancer Immunol. Res.* **5**, 898–907 (2017).
10. Kessel, A. et al. *Autoimmun. Rev.* **11**, 670–677 (2012).
11. Khan, A. R. et al. *Nature Commun.* **6**, 5997 (2015).
12. Sautès-Fridman, C., Petitprez, F., Calderaro, J. & Fridman, W. H. *Nature Rev. Cancer* **19**, 307–325 (2019).
13. Affara, N. I. et al. *Cancer Cell* **25**, 809–821 (2014).
14. Shalapour, S. et al. *Nature* **551**, 340–345 (2017).
15. Ammirante, M. et al. *Nature* **464**, 302–305 (2010).

This article was published online on 15 January 2020.



WALTER STEIN/GETTY

## Invasive plants versus herbivores

The population of large animals in the Gorongosa National Park collapsed during the Mozambican civil war (1977–92), and led to encroachment of the invasive shrub *Mimosa pigra*. Writing in *Nature Ecology & Evolution*, Guyton et al. report that Gorongosa's repopulation with large herbivores has reduced the abundance of mimosa to pre-war levels (J. A. Guyton et al. *Nature Ecol. Evol.* <http://doi.org/djff>; 2020).

By analysing faecal samples from Gorongosa's five main ruminant herbivores, including waterbuck (*Kobus ellipsiprymnus*; pictured), the authors found that mimosa was the main component of the diets of these

species in 2013–18. They also found that the shrub's density and biomass were greater in fenced enclosures that excluded herbivores than in unfenced areas.

The authors therefore conclude that the burgeoning populations of native large herbivores are consuming mimosa, and have thereby conferred resistance to its invasion in just ten years. The findings suggest that rewilding is a potentially useful strategy for reversing a common form of environmental degradation in Africa's protected areas. **Andrew Mitchinson**

This article was published online on 16 January 2020.



# Versatile strategy for making 2D materials

Wei Sun Leong

Two-dimensional materials have potential uses in flexible electronics, biosensors and water purification. A method for producing air-stable 2D materials on an industrial scale, now reported, is a key step in bringing them to market. **See p.492**

Modern materials science relies on a deep understanding of defects – interruptions to regular atomic arrangements in crystalline solids. Although ‘defects’ brings to mind imperfections and blemishes, they often make a material more useful than it otherwise would be. For example, metal impurities such as chromium and iron atoms in corundum (a crystalline form of aluminium oxide) are responsible for the colours of rubies and sapphires. Moreover, the addition of impurities to silicon has enabled the current era of computing and robotics. On page 492, Du *et al.*<sup>1</sup> report a method for producing a variety of technologically useful two-dimensional materials that contain deliberately introduced impurities, solving a fabrication problem for next-generation devices.

Transition-metal chalcogenides (TMCs) are emerging materials that hold great promise for their incorporation into a wide range of applications, from batteries and flexible electronics to biosensors and water-purification systems. They are composed of a transition metal such as molybdenum or tungsten and a chalcogen (an element in group 16 of the periodic table) such as sulfur, selenium or tellurium. The properties of TMC monolayers change greatly if the metallic element is altered. In particular, these structures can change from being normal metals to semiconductors, or even superconductors.

In the past few years, many researchers<sup>2–4</sup> have focused on making ultrathin electronics that have superior properties to those of existing silicon devices, by combining different TMC monolayers into a single object known as a heterostructure, using a technique called chemical-vapour deposition. Other researchers<sup>5</sup> have produced functional devices using a single TMC in which different regions of the material have different properties, such as being metallic or semiconducting. However, although these techniques are good for fabricating prototype devices, they are not practical enough for real-world applications.

The long-standing problem in incorporating

TMC monolayers into a functional device has been the lack of a metallic-phase TMC monolayer that is stable in ambient conditions for more than a month<sup>6</sup>. Du and colleagues overcame this challenge, and made metallic-phase TMC monolayers that they show can exist in such conditions for about a year. The authors achieved this feat by introducing a technology based on a process known as doping.

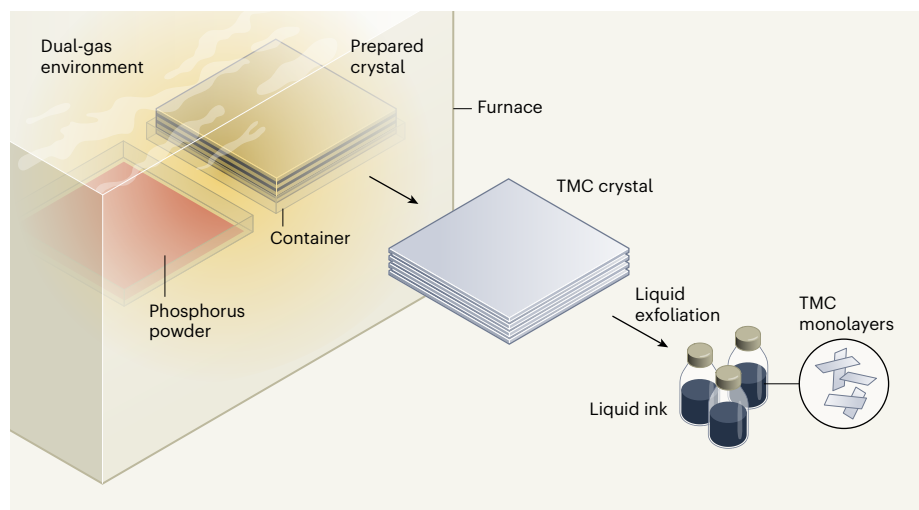
Doping has shaped the digital revolution – the shift from analog to digital electronics that began in the second half of the twentieth century. The process involves changing the electrical conductivities of semiconductors such as silicon by adding impurities. Eighty years ago<sup>7</sup>, dopant atoms of boron and phosphorus were added to pure silicon to produce materials called p-type and n-type silicon, respectively; these form p–n junctions, the basis of computing. This doping technology

continues to be useful today, and is found in our everyday electronics. Du and co-workers’ doping technology for 2D materials is also expected to have a long-term impact on the field.

The authors produced TMC monolayers in three steps (Fig. 1). First, they prepared a crystal that contained two different transition metals (one of which provided impurity atoms for TMC doping), an element in group 13 or 14 of the periodic table, and carbon. Second, they heated the crystal at high temperatures (873–1,373 kelvin) for 4 hours in an environment that contained two gases. One of these was a chalcogen-containing gas that supplied chalcogen atoms for the TMC; the other gas was phosphorus, which provided further impurity atoms for TMC doping. Third, the authors used a process called liquid exfoliation to convert the resulting TMC crystal into TMC monolayers in the form of liquid inks.

Du *et al.* used this three-step dual-doping technology to make, for example, metallic-phase TMC monolayers of tungsten disulfide that were doped with both yttrium and phosphorus atoms. They also produced undoped TMC monolayers by preparing layered crystals that contained one type of transition metal, rather than two, and removing the source of phosphorus gas. In total, the authors made six doped and seven undoped TMC monolayers, demonstrating the remarkable versatility of their approach for producing 2D materials.

One major advantage of Du and colleagues’ method is that the final 2D materials are in the form of liquid inks. There is clearly a shift in this field towards making high-quality monolayer



**Figure 1 | Method for producing air-stable transition-metal chalcogenides (TMCs).** Du *et al.*<sup>1</sup> demonstrate a technology for making monolayers of materials called TMCs that they show can remain stable in ambient conditions for about a year. They first prepare a crystal that contains two different transition metals, an element in group 13 or 14 of the periodic table, and carbon. They then place the crystal in a container and heat it in a furnace for 4 hours, in an environment containing two gases. One of the gases contains a chalcogen (an element in group 16 of the periodic table) and the other is phosphorus gas produced by heating phosphorus powder in a separate container in the furnace. The result of this process is a TMC crystal. Finally, the authors use a process called liquid exfoliation to convert the crystal into TMC monolayers in the form of liquid inks.

inks for commercialization<sup>8,9</sup>, rather than films produced by techniques such as epitaxial growth or chemical-vapour deposition. Such films require a process known as delamination to separate them from their growth substrates, which deteriorates the material's quality and necessitates further processing<sup>10,11</sup>. By contrast, monolayer inks can be readily deposited on arbitrary substrates using techniques such as inkjet printing or spin coating, and so are easily integrated into 3D systems<sup>12,13</sup>.

From a scientific standpoint, 2D materials need to be stable and usable in our immediate surroundings. Du and colleagues' findings are promising for the field because they show that the presence of a low quantity (less than 1%) of impurity atoms can stabilize TMC monolayers. This result suggests that materials researchers should start to explore the use of chemical elements to stabilize 2D materials that would otherwise degrade in ambient conditions within hours, rather than using encapsulation layers, which complicate the monolayer systems.

The next steps will be for theorists to predict suitable 'impurity stabilizers' for TMC monolayers, and for experimentalists to investigate the use of elements that are abundant on Earth. In the meantime, it should still be possible to build advanced machines for precise and reliable dual doping of TMCs, because only a low quantity of relatively rare yttrium and phosphorus is needed to stabilize TMC monolayers. Du and colleagues' work demonstrates that, whatever new materials are discovered, it is crucial that we understand, manipulate and use their atomic-level defects. Every atom matters.

**Wei Sun Leong** is in the Department of Materials Science and Engineering, National University of Singapore, 117575 Singapore. e-mail: weisun@u.nus.edu

## Microbiology

# Meet the relatives of our cellular ancestor

Christa Schleper & Filipa L. Sousa

Microorganisms related to lineages of the Asgard archaea group are thought to have evolved into complex eukaryotic cells. Now the first Asgard archaeal species to be grown in the laboratory reveals its metabolism and cell biology. **See p.519**

Complex life forms including plants, animals and fungi are known as eukaryotes. These organisms are composed of cells that contain membrane-bound internal compartments such as nuclei and other organelles. Imachi *et al.*<sup>1</sup> report on page 519 that a type of microorganism called an Asgard archaeon, which might shed light on how early eukaryotic cells evolved, has finally been cultured in the laboratory. The achievement will enable detailed metabolic and cellular investigation of microbes that represent the closest Archaeal relative of eukaryotes cultured so far.

It is thought that eukaryotes arose when two types of single cell merged, with one engulfing the other. A cell from the domain archaea is proposed to have engulfed a bacterial cell of a type known as an alphaproteobacterium, and the engulfed bacterium evolved into eukaryotes' energy-generating organelles – mitochondria.

However, the nature of the ancestral cell that engulfed this bacterium is unclear. Genomic analyses have strengthened the idea that this cell traces back to archaea because many archaeal genes involved in central biological processes such as transcription, translation and DNA replication share a common ancestry with (are phylogenetically related to) the corresponding eukaryotic genes. Was the alphaproteobacterium engulfed by a bona fide archaeal cell, or by an archaeal cell that had already acquired some eukaryotic characteristics, such as a nucleus? No fossils have been found that could shed light on the early eukaryotic ancestors. However, investigation of archaeal lineages has offered a way forward.

Since 2015, on the basis of genomic and phylogenetic analyses<sup>2</sup>, archaea of a newly discovered phylum termed Lokiarchaeota (after the Norse god Loki) have been proposed as the closest living relatives of the ancient archaeal host cells from which eukaryotes are thought to have evolved. Subsequent genomic research revealed yet more such lineages, for which other Norse gods have provided names (Thor, Odin, Heimdall and Hel)<sup>3,4</sup>, and which are now

grouped together with Lokiarchaeota into what are collectively termed Asgard archaea (Fig. 1). Intriguingly, all of these lineages contain an unprecedentedly large number of genes that encode what are called eukaryotic signature proteins (ESPs), which are usually found only in eukaryotes<sup>2,3,5,6</sup>. Heimdallarchaeota currently represent the predicted closest Archaeal relative of eukaryotes on the basis of phylogenetic analysis and the ESP content of their genomes<sup>3,7</sup>. However, all members of the Asgard archaea were previously identified, and their metabolism predicted, solely by their DNA sequences, and thus their cellular features have remained unknown until now.

Imachi and colleagues report that they have cultured in the laboratory an Asgard archaeon from the Lokiarchaeota phylum that they propose to call 'Prometheoarchaeum syntrophicum', which was obtained from deep-ocean sediments. The unusual shape and metabolism of Prometheoarchaeum prompt the authors to propose a new model for the emergence of the first eukaryotic cell. This event, predicted<sup>8</sup> to have occurred between 2 billion and 1.8 billion years ago, is one of the key cellular transitions in evolutionary biology, and is also a major biological mystery.

More than six years before Asgards were even identified, Imachi and colleagues had already started to generate enrichment cultures of microorganisms found in deep marine sediments<sup>9</sup>. Their original goal was to find organisms that could degrade methane, and the authors searched for such microbes at a site about 2.5 kilometres below the ocean surface off the coast of Japan.

Imachi *et al.* set up a flow bioreactor device that mimicked the temperature (10 °C) and the low-oxygen and low-nutrient conditions at this underwater site. Within five years of starting this bioreactor work, a highly diverse consortium of active bacteria and archaea, including Lokiarchaeota, were obtained. Small subcultures were then used to gradually enrich for cultures in which archaeal cells were the dominant component, and

1. Du, Z. *et al.* *Nature* **577**, 492–496 (2020).
2. Sahoo, P. K., Memaran, S., Xin, Y., Balicas, L. & Gutiérrez, H. R. *Nature* **553**, 63–67 (2018).
3. Leong, W. S. *et al.* *J. Am. Chem. Soc.* **140**, 12354–12358 (2018).
4. Zhang, Z. *et al.* *Science* **357**, 788–792 (2017).
5. Kappera, R. *et al.* *Nature Mater.* **13**, 1128–1134 (2014).
6. Lin, H. *et al.* *Nature Mater.* **18**, 602–607 (2019).
7. Hey, T. & Pápay, G. in *The Computing Universe: A Journey through a Revolution 123* (Cambridge Univ. Press, 2014).
8. Lin, Z. *et al.* *Nature* **562**, 254–258 (2018).
9. Pan, K. *et al.* *Nature Commun.* **9**, 5197 (2018).
10. Shim, J. *et al.* *Science* **362**, 665–670 (2018).
11. Leong, W. S. *et al.* *Nature Commun.* **10**, 867 (2019).
12. McManus, D. *et al.* *Nature Nanotechnol.* **12**, 343–350 (2017).
13. Sivan, M. *et al.* *Nature Commun.* **10**, 5201 (2019).



# News & views

## Microbiology

# Meet the relatives of our cellular ancestor

Christa Schleper & Filipa L. Sousa

Microorganisms related to lineages of the Asgard archaea group are thought to have evolved into complex eukaryotic cells. Now the first Asgard archaeal species to be grown in the laboratory reveals its metabolism and cell biology.

Complex life forms including plants, animals and fungi are known as eukaryotes. These organisms are composed of cells that contain membrane-bound internal compartments such as nuclei and other organelles. Writing in *Nature*, Imachi *et al.*<sup>1</sup> report that a type of microorganism called an Asgard archaeon, which might shed light on how early eukaryotic cells evolved, has finally been cultured in the laboratory. The achievement will enable detailed metabolic and cellular investigation of microbes that represent the closest Archaeal relative of eukaryotes cultured so far.

It is thought that eukaryotes arose when two types of single cell merged, with one engulfing the other. A cell from the domain archaea is proposed to have engulfed a bacterial cell of a type known as an alphaproteobacterium, and the engulfed bacterium evolved into eukaryotes' energy-generating organelles – mitochondria.

However, the nature of the ancestral cell that engulfed this bacterium is unclear. Genomic analyses have strengthened the idea that this cell traces back to archaea because many archaeal genes involved in central biological processes such as transcription, translation and DNA replication share a common ancestry with (are phylogenetically related to) the corresponding eukaryotic genes. Was the alphaproteobacterium engulfed by a bona fide archaeal cell, or by an archaeal cell that had already acquired some eukaryotic characteristics, such as a nucleus? No fossils have been found that could shed light on the early eukaryotic ancestors. However, investigation of archaeal lineages has offered a way forward.

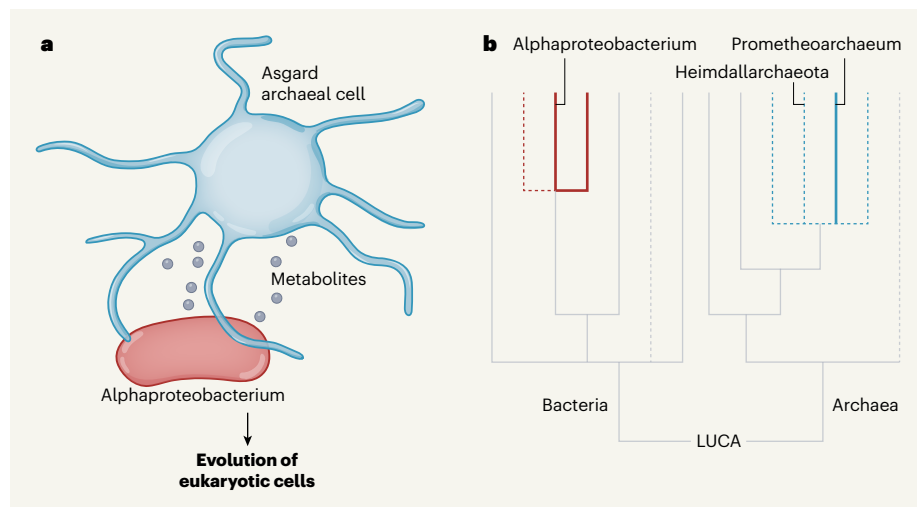
Since 2015, on the basis of genomic and phylogenetic analyses<sup>2</sup>, archaea of a newly discovered phylum termed Lokiarchaeota (after the Norse god Loki) have been proposed as the

closest living relatives of the ancient archaeal host cells from which eukaryotes are thought to have evolved. Subsequent genomic research revealed yet more such lineages, for which other Norse gods have provided names (Thor, Odin, Heimdall and Hel)<sup>3,4</sup>, and which are now grouped together with Lokiarchaeota into what are collectively termed Asgard archaea (Fig. 1). Intriguingly, all of these lineages

contain an unprecedentedly large number of genes that encode what are called eukaryotic signature proteins (ESPs), which are usually found only in eukaryotes<sup>2,3,5,6</sup>. Heimdallarchaeota currently represent the predicted closest Archaeal relative of eukaryotes on the basis of phylogenetic analysis and the ESP content of their genomes<sup>3,7</sup>. However, all members of the Asgard archaea were previously identified, and their metabolism predicted, solely by their DNA sequences, and thus their cellular features have remained unknown until now.

Imachi and colleagues report that they have cultured in the laboratory an Asgard archaeon from the Lokiarchaeota phylum that they propose to call 'Prometheoarchaeum syntrophicum', which was obtained from deep-ocean sediments. The unusual shape and metabolism of Prometheoarchaeum prompt the authors to propose a new model for the emergence of the first eukaryotic cell. This event, predicted<sup>8</sup> to have occurred between 2 billion and 1.8 billion years ago, is one of the key cellular transitions in evolutionary biology, and is also a major biological mystery.

More than six years before Asgards were even identified, Imachi and colleagues had



**Figure 1 | The evolution of eukaryotic cells.** Imachi *et al.*<sup>1</sup> report that they have cultured a microorganism, which they call 'Prometheoarchaeum syntrophicum', in the laboratory. The microbe belongs to a group known as Asgard archaea. This is the first time that an Asgard archaeon has been cultured, and has revealed previously unknown aspects of its cellular biology, including the presence of long protrusions. This development might shed light on how complex eukaryotic cells evolved. **a**, It is thought that an ancient Asgard archaeon interacted with a bacterium from the class Alphaproteobacteria, for example by exchanging metabolite molecules (grey circles). The mitochondrion, the energy-generating organelle of eukaryote cells, is thought to have evolved when such a bacterium was taken up in the archaeal cell. **b**, This simplified evolutionary tree includes branches of the lineages (Proteobacteria shown in red and Asgard archaea in blue) that might have contributed to the formation of eukaryotic cells. Dashed lines on the evolutionary trees represent lineages identified only by genomic analysis and not by organisms cultured in the laboratory. It is thought that eukaryotic cells evolved from a partnership between an alphaproteobacterium and a relative of a Heimdallarchaeote (neither of which is known). LUCA: the last universal common ancestor (the cell(s) from which bacteria and archaea evolved).

already started to generate enrichment cultures of microorganisms found in deep marine sediments<sup>9</sup>. Their original goal was to find organisms that could degrade methane, and the authors searched for such microbes at a site about 2.5 kilometres below the ocean surface off the coast of Japan.

Imachi *et al.* set up a flow bioreactor device that mimicked the temperature (10 °C) and the low-oxygen and low-nutrient conditions at this underwater site. Within five years of starting this bioreactor work, a highly diverse consortium of active bacteria and archaea, including Lokiarchaeota, were obtained. Small subcultures were then used to gradually enrich for cultures in which archaeal cells were the dominant component, and Prometheoarchaeum was successfully enriched in this way after seven more years of work. These optimizations revealed that Prometheoarchaeum grows best in conditions that do not directly reflect its original habitat: at 20 °C and supplemented with amino acids, peptides and even baby-milk powder.

The authors report that Prometheoarchaeum's growth depends on the presence of other microbial partners that in turn rely on Prometheoarchaeum for their survival – a relationship called a syntrophy. The partners scavenge hydrogen released by Prometheoarchaeum, a metabolic product that was correctly predicted to be generated by Asgard archaea on the basis of genomic data<sup>5</sup>. The authors found that Prometheoarchaeum could be enriched to make up more than 80% of the cells in the culture, even though it grows extremely slowly, taking 2 to 4 weeks to replicate and divide. From preliminary studies using isotope analysis, the authors report that this organism can degrade externally supplied amino acids. However, that does not exclude the possibility that it also thrives on other nutrients in the growth medium.

Prometheoarchaeum cells are relatively small (300–750 nanometres in diameter), have lipids characteristic of other archaea, and show no evidence for eukaryotic-like organelles. However, the organism forms intriguing

structures on its cellular surface that include long and often branching protrusions.

On the basis of its cell shape and small size, and on evidence that Prometheoarchaeum produces and syntrophically transfers hydrogen and formate molecules to other organisms, the authors propose a new model for the emergence of eukaryotic cells – one involving three partners. In this model, a free-living bacterial ancestor that would give rise to mitochondria became entangled with, and was then engulfed by, an archaeal host cell that itself was in a syntrophic relationship with a bacterial partner.

This model is consistent with earlier suggestions about the engulfment process in eukaryotic evolution<sup>10</sup>, and emphasizes the importance of membrane-mediated processes in the origin of eukaryotes<sup>11</sup>. However, extensive cellular protrusions are not found exclusively in this Asgard archaeon. It would therefore be of interest to investigate to what extent these protrusions differ from those of branched cellular extensions previously observed in other archaea such as *Pyrodictum*<sup>12</sup> or *Thermococcus* species<sup>13</sup>. In addition, it will be interesting to determine whether the ESPs potentially involved in membrane remodelling are localized in these structures in Prometheoarchaeum.

The syntrophic interactions that Imachi and colleagues propose in their model for the origin of mitochondria are based on the need for the host cell to adapt to oxygen use (as a consequence of rising oxygen levels on the ancient Earth). These ideas differ from the 'reverse hydrogen flow' model, which suggests instead that hydrogen produced by the archaeon is consumed directly by the bacterial mitochondrial ancestor, with no need to invoke a hypothetical third partner<sup>5</sup>. Considering that Prometheoarchaeum does not directly represent the archaeal ancestor of eukaryotes (nor does any other currently existing archaeon), other suggested metabolic exchanges between the archaeal host and bacterial mitochondrial ancestor, such as hydrogen consumption from the archaeal<sup>14,15</sup>

or the bacterial side<sup>5</sup>, remain plausible as initial drivers of a syntrophic relationship. In any case, the many models for the origin of eukaryotes<sup>5,11,14,15</sup> highlight the importance of initial syntrophic associations<sup>5,14,15</sup> and membrane-mediated processes<sup>10,11</sup>. Interestingly, albeit for different reasons, both syntrophy and membranes were crucial aspects in an engineered synthetic relationship in which an *Escherichia coli* bacterium was maintained inside a yeast cell for more than 120 days<sup>16</sup>.

Imachi and colleagues' success in culturing Prometheoarchaeum after efforts spanning more than a decade represents a huge breakthrough for microbiology. It sets the stage for the use of molecular and imaging techniques to further elucidate the metabolism of Prometheoarchaeum and the role of ESPs in archaeal cell biology. This, in turn, could guide the direction of future work investigating how eukaryotic cells emerged.

**Christa Schleper** and **Filipa L. Sousa** are in the Archaea Biology and Ecogenomics Division, University of Vienna, 1080 Vienna, Austria.  
e-mails: christa.schleper@univie.ac.at;  
filipa.sousa@univie.ac.at

1. Imachi, H. *et al.* *Nature* <https://doi.org/10.1038/s41586-019-1916-6> (2020).
2. Spang, A. *et al.* *Nature* **521**, 173–179 (2015).
3. Zaremba-Niedzwiedzka, K. *et al.* *Nature* **541**, 353–358 (2017).
4. Seitz, K. W. *et al.* *Nature Commun.* **10**, 1822 (2019).
5. Spang, A. *et al.* *Nature Microbiol.* **4**, 1138–1148 (2019).
6. Caspi, Y. & Dekker, C. *Front. Microbiol.* **9**, 174 (2018).
7. Williams, T. A., Cox, C. J., Foster, P. G., Szöllősi, G. J. & Embley, T. M. *Nature Ecol. Evol.* **4**, 138–147 (2020).
8. Betts, H. C. *et al.* *Nature Ecol. Evol.* **2**, 1556–1562 (2018).
9. Imachi, H. *et al.* *ISME J.* **5**, 1913–1925 (2011).
10. Baum, D. A. & Baum, B. *BMC Biol.* **12**, 76 (2014).
11. Gould, S. B., Garg, S. G. & Martin, W. F. *Trends Microbiol.* **24**, 525–534 (2016).
12. Rieger, G., Rachel, R., Hermann, R. & Stetter, K. O. *J. Struct. Biol.* **115**, 78–87 (1995).
13. Miroshnichenko, M. L. *et al.* *Int. J. Syst. Bacteriol.* **48**, 23–29 (1998).
14. Martin, W. & Müller, M. *Nature* **392**, 37–41 (1998).
15. López-García, P. & Moreira, D. *Trends Biochem. Sci.* **24**, 88–93 (1999).
16. Mehta, A. P. *et al.* *Proc. Natl Acad. Sci. USA* **115**, 11796–11801 (2018).



# Signatures of self-organized criticality in an ultracold atomic gas

<https://doi.org/10.1038/s41586-019-1908-6>

Received: 25 June 2018

Accepted: 23 October 2019

Published online: 15 January 2020

S. Helmrich<sup>1</sup>, A. Arias<sup>1,2,3</sup>, G. Lochead<sup>1,2,3</sup>, T. M. Wintermantel<sup>1,2,3</sup>, M. Buchhold<sup>4,5</sup>, S. Diehl<sup>6</sup> & S. Whitlock<sup>1,2,3\*</sup>

Self-organized criticality is an elegant explanation of how complex structures emerge and persist throughout nature<sup>1</sup>, and why such structures often exhibit similar scale-invariant properties<sup>2–9</sup>. Although self-organized criticality is sometimes captured by simple models that feature a critical point as an attractor for the dynamics<sup>10–15</sup>, the connection to real-world systems is exceptionally hard to test quantitatively<sup>16–21</sup>. Here we observe three key signatures of self-organized criticality in the dynamics of a driven–dissipative gas of ultracold potassium atoms: self-organization to a stationary state that is largely independent of the initial conditions; scale-invariance of the final density characterized by a unique scaling function; and large fluctuations of the number of excited atoms (avalanches) obeying a characteristic power-law distribution. This work establishes a well-controlled platform for investigating self-organization phenomena and non-equilibrium criticality, with experimental access to the underlying microscopic details of the system.

Self-organized criticality (SOC) was conceptualized as a way to explain the abundance of scale-invariant systems found in nature<sup>1</sup>. It is thought to underlie a range of complex dynamical phenomena, from activity in electrical circuits and neural networks<sup>7,9</sup>, to the likelihood of avalanches and earthquakes<sup>2</sup> as well as how forest fires<sup>4,10</sup>, diseases<sup>3</sup> and even ideas spread<sup>8</sup>. However, despite the fundamental and practical importance of SOC phenomena, much-needed controlled experiments are hindered by numerous complexities concerning the relevant microscopic degrees of freedom<sup>16–19</sup> and even the simplest models (beyond mean-field approximations) present serious challenges to theory<sup>11–15,20</sup>.

SOC can be understood as an organizing principle that governs a class of dissipative interacting systems that display three key signatures: (1) self-organization to a stationary state (bringing observables to values that are independent of initial conditions); (2) scale invariance of spatio-temporal correlation functions, including bulk observables; and (3) critical response to small perturbations, usually encountered in the form of avalanches that have a broad range of sizes and durations and that are described by power-law distributions. This differs from an equilibrium phase transition, where scale invariance and a critical response ensue only for a fine-tuned parameter set. The common root of these emergent SOC properties is that the respective gap (that is, the distance in parameter space from the critical state) is replaced by a ‘dynamical gap’ that self-tunes to zero by an intrinsic feedback mechanism. This property, and signatures (1)–(3), set SOC apart from other occurrences of non-equilibrium scaling behaviour—such as hydrodynamic long-time tails<sup>22</sup>, the Kosterlitz–Thouless critical phase in two-dimensional quantum fluids<sup>23</sup> and the transient dynamics of turbulent cascades in isolated systems<sup>24,25</sup>—which have also been studied with ultracold atoms<sup>26–31</sup>; see also related experiments on superradiance<sup>32,33</sup> and scaling in unitary Bose gases<sup>34</sup>.

In this work, we demonstrate signatures (1)–(3) of SOC in a microscopically well-controlled physical system: a three-dimensional trapped gas of ultracold potassium atoms driven to highly excited Rydberg states by a laser field (Fig. 1a). The ingredient that leads to SOC is the slow, irreversible decay of the excited population to auxiliary inactive states, which has been largely disregarded in the investigation of Rydberg many-body dynamics. This enables the observation of a phase transition from a self-organizing active phase to an absorbing phase, scale-invariance of the self-organized density and large fluctuations of the active density in the form of power-law distributed avalanches. Beyond these experimental results, we derive a Langevin equation from the underlying microscopic many-body quantum master equation that governs driven–dissipative Rydberg dynamics, which coincides with one of the emblematic classes of SOC models<sup>20</sup>. This provides the crucial link from the microscopic atomic physics to the observed macroscopic SOC phenomenology, and establishes ultracold Rydberg atomic gases as a widely tunable and theoretically accessible platform for studying self-organization and universality in non-equilibrium dynamics.

## Physical system

Each of the approximately  $10^5$  atoms held in the optical trap can be represented by a three state system: a ground state  $|g\rangle = |4s_{1/2}, F=1\rangle$ , an excited Rydberg state  $|r\rangle$ , and auxiliary removed states, which we refer to collectively by  $|0\rangle$  (Fig. 1b). The laser field drives the  $|g\rangle \rightarrow |r\rangle$  transition with a fixed detuning  $\Delta$  from resonance and with an amplitude parameterized by the Rabi frequency  $\Omega$ . In our experiments  $\Delta \gg \Omega$ , such that single-atom excitation processes are strongly suppressed. Once excited, however, atoms can facilitate further excitations (when the laser detuning is compensated by the interaction energy of Rydberg

<sup>1</sup>Physikalisches Institut, Universität Heidelberg, Heidelberg, Germany. <sup>2</sup>Institut de Science et d'Ingénierie Supramoléculaires (ISIS, UMR 7006), University of Strasbourg and CNRS, Strasbourg, France. <sup>3</sup>Institut de Physique et Chimie des Matériaux de Strasbourg (IPCMS, UMR 7504), University of Strasbourg and CNRS, Strasbourg, France. <sup>4</sup>Department of Physics, California Institute of Technology, Pasadena, CA, USA. <sup>5</sup>Institute for Quantum Information and Matter, California Institute of Technology, Pasadena, CA, USA. <sup>6</sup>Institut für Theoretische Physik, Universität zu Köln, Cologne, Germany. \*e-mail: whitlock@unistra.fr

pair states) leading to the formation of extended excitation clusters<sup>35–44</sup>. Alternatively, they can be lost from the system, predominantly by spontaneously decaying to another hyperfine ground state or to other untrapped states.

This system permits a microscopic description via a quantum master equation for the many-body density matrix  $\hat{\rho}$ ,

$$\partial_t \hat{\rho} = \frac{i}{\hbar} [\hat{\rho}, \hat{H}] + \sum_l \mathcal{L}_l(\hat{\rho}) \quad (1)$$

with the atom–light interaction Hamiltonian  $\hat{H}$  and Lindblad superoperator  $\mathcal{L}_l(\hat{\rho})$  given by

$$\hat{H} = \sum_l \left[ \left( \sum_{l'} \frac{1}{2} \frac{C_6}{|\mathbf{r}_{ll'}|^6} \hat{\sigma}_l^{rr} - \Delta \right) \hat{\sigma}_l^{rr} + \frac{\Omega}{2} (\hat{\sigma}_l^{gr} + \hat{\sigma}_l^{rg}) \right]$$

$$\mathcal{L}_l(\hat{\rho}) = \Gamma \hat{\sigma}_l^{or} \hat{\rho} \hat{\sigma}_l^{ro} + \gamma_{de} \hat{\sigma}_l^{rr} \hat{\rho} \hat{\sigma}_l^{rr} - \frac{\gamma_{de} + \Gamma}{2} (\hat{\sigma}_l^{rr} \hat{\rho} + \hat{\rho} \hat{\sigma}_l^{rr})$$

where  $\hat{\sigma}_l^{ab} \equiv |\alpha\rangle\langle\beta|_l$ ,  $l, l'$  are indices for each atom and  $\mathbf{r}_{ll'}$  is the relative distance between any two atoms. Interactions between Rydberg states are parameterized by the van der Waals coefficients  $C_6/2\pi \approx 0.52 \text{ GHz } \mu\text{m}^6$  for  $|r\rangle = |39p_{3/2}\rangle$ , and  $C_6/2\pi \approx 238 \text{ GHz } \mu\text{m}^6$  for  $|r\rangle = |66p_{3/2}\rangle$ . Dissipation is described by  $\mathcal{L}_l(\hat{\rho})$ , which includes decay (with total rate  $\Gamma$ ) and dephasing (rate  $\gamma_{de}$ ) attributed primarily to residual laser phase noise and Doppler broadening.

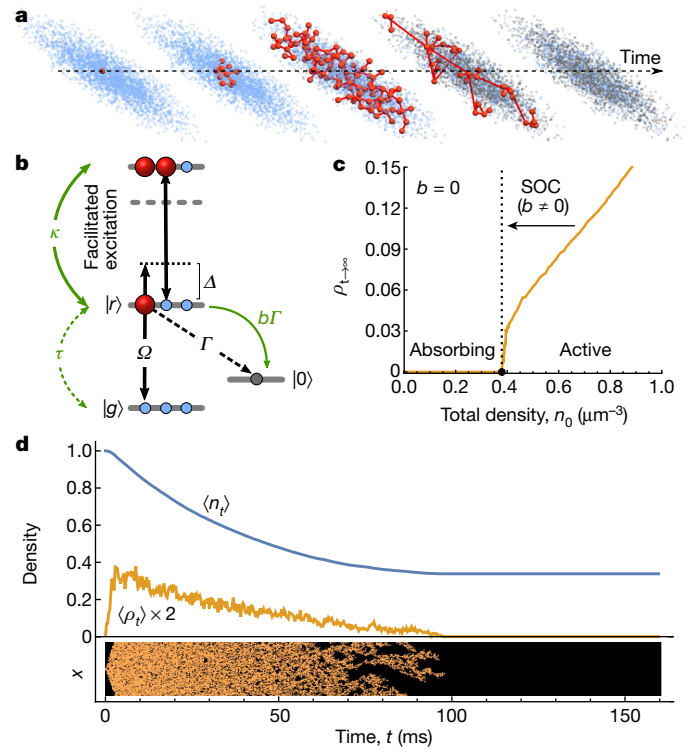
To connect the microscopic dynamics of Rydberg atoms (equation (1)) to the SOC phenomenology, we apply a systematic coarse graining procedure for the collective dynamics (derived in the Supplementary Information). In brief, we average over the characteristic length scale of the facilitation process and adiabatically eliminate the rapidly decaying atomic coherences<sup>45,46</sup>. We also approximate the atomic medium as a quasi-homogeneous gas with a smoothly varying density, which is justified by the fact that the atoms move on a timescale considerably shorter than the SOC dynamics. The final result is a Langevin equation for the spatio-temporal density of atoms in the  $|r\rangle$  state,  $\rho_r = \rho(t, \mathbf{r})$ , (the active component) and the total remaining density  $n_t = n(t, \mathbf{r})$ , which is the sum of the populations in the  $|g\rangle$  and  $|r\rangle$  states (excluding removed states):

$$\partial_t \rho_r = (D\nabla^2 - \Gamma + \kappa n_t) \rho_r - 2\kappa \rho_r^2 + \tau(n_t - 2\rho_r) + \xi_t \quad (2)$$

$$n_t = n_0 - b\Gamma \int_0^t dt' \rho_{r,t'} + D_T \int_0^t dt' \nabla^2 n_t \quad (3)$$

In these equations,  $D$  and  $D_T$  are diffusion constants and  $\kappa$  is the facilitation rate (which together govern the rate of excitation spreading),  $\tau$  is the spontaneous excitation rate,  $n_0$  is the initial density, and  $b$  is a dimensionless parameter that governs how fast the decay depletes the total population. The stochastic part of the evolution is governed by the autocorrelated multiplicative noise term  $\xi_t = \xi(t, \mathbf{r})$  with variance  $\text{var}(\xi_t) = \Gamma \rho_r$ .

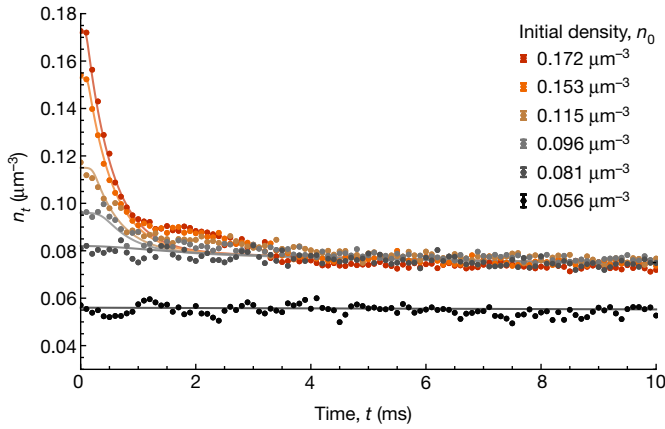
Equations (2) and (3) closely relate to the paradigmatic Drossel–Schwabl forest fire model<sup>10,20</sup>, except for the absence of a slow regrowth term for the total density, which would normally bring the system from an inactive (subcritical) state to the critical state. This regrowth is typically the slowest scale in the model, and must asymptotically vanish in order to realize SOC. Nevertheless, in its absence the system still exhibits a non-equilibrium phase transition<sup>20</sup>, which can be approached by starting in the active phase. To illustrate this we present numerical simulations (Fig. 1c, d), for simplicity focusing on a small one-dimensional system. In the case  $b = \tau = 0$ , the system features a non-equilibrium phase transition<sup>44</sup> from an absorbing phase, in which any excited component quickly dies out (characterized by  $\rho_{r,\infty} \rightarrow 0$  for  $\kappa n_0 \ll \Gamma$ ), to



**Fig. 1 | SOC in an ultracold atomic gas excited to Rydberg states by a laser field.** **a**, Self-organization process in a cigar-shaped atom cloud showing atoms in the ground state  $|g\rangle$  (blue dots) or excited to a Rydberg state  $|r\rangle$  (large red spheres) via facilitated excitation processes, leading to the buildup of correlations (represented by red links). **b**, The laser field couples the  $|g\rangle \rightarrow |r\rangle$  transition with Rabi frequency  $\Omega$  and detuning  $\Delta$ , and atoms in the  $|r\rangle$  state either decay to removed states  $|0\rangle$  (black circles) or facilitate further Rydberg excitations. These microscopic processes determine the couplings in the Langevin equation (equations (2) and (3)) defined in the text (green arrows). **c**, Numerical solution of equations (2) and (3) for the population conserving system  $b = 0$  (in one dimension) with  $D = 1$  (discretization distance = 1),  $D_T = 0$ ,  $\Gamma = 10$ ,  $\kappa = 10$  and  $\tau = 0$ . As a function of the total density  $n_0$ , the stationary active density  $\rho_{r,\infty}$  exhibits an absorbing state phase transition (dotted vertical line), which acts as an attractor for the SOC dynamics (when  $b \neq 0$ ). **d**, Time evolution for  $b = 0.01$  showing the spatially averaged active density  $\langle \rho_r \rangle$  (orange) and total density  $\langle n_t \rangle$  (blue) as the system approaches a stationary state close to the critical point of the absorbing state phase transition. The lower panel in **d** shows the full spatio-temporal evolution of the active density  $\rho_r$  with transverse coordinate  $x$  spanning 128 grid points.

an active phase in which excitations spread throughout the system from arbitrarily small seed excitations (with  $\rho_{r,t=0} > 0$  for  $\kappa n_0 \gg \Gamma$ ). On the other hand, when  $b, \tau \neq 0$ , spontaneous single-atom excitations trigger the relatively fast facilitated excitation dynamics, although on longer timescales particle loss introduces a coupling between  $\rho_r$  and  $n_t$ . Specifically, the first integral in equation (3) acts as a feedback mechanism, causing  $n_t$  to continuously decrease while in the active phase. When this loss is much slower than the internal dynamics but much faster than the spontaneous excitation rate (achieved for  $\kappa n_0 \approx \Gamma \gg b\Gamma \gg \tau$ ), the system slowly approaches the critical point of the absorbing-state phase transition and develops scale-invariant properties, visualized for example by growing spatio-temporal correlations in the active density (the fractal-like structures seen around  $t = 80$  ms in the lower panel of Fig. 1d). This behaviour can be understood in terms of the evolution of the dynamical gap  $\kappa n_t - \Gamma$ , which is initially positive and then continuously decreases—owing to population loss—until it asymptotically reaches zero at the critical point, where the dynamics effectively stop.





**Fig. 2 | Self-organization: above a threshold value, the remaining total atom density  $n_t$  is attracted to the same stationary-state density independent of the initial conditions.** The Rydberg state used is  $39p_{3/2}$  and the parameters of the driving laser field are  $\Delta/2\pi = 30$  MHz,  $\Omega/2\pi = 190$  kHz. For high initial densities  $n_0 \geq 0.08 \mu\text{m}^{-3}$  the time dependence consists of a short initial plateau followed by fast exponential decay to a stationary state with a fixed density  $n_t = 0.075 \mu\text{m}^{-3}$ . For initial densities below  $n_t$  the dynamics is effectively stationary (black points). The solid lines correspond to mean-field solutions to the effective Langevin equation with parameters given in the text. Each data point is the average of three measurements. Standard errors for each dataset are indicated by the representative error bars shown in the key.

Theoretically solving the full dynamical many-body problem described by equations (2) and (3) beyond the mean-field level is difficult, particularly for large system sizes and in more than one spatial dimension, owing to the presence of multiplicative noise and the importance of strong spatio-temporal correlations<sup>11,14</sup>. As a result, many properties of this class of systems are still actively debated, such as the question of whether the system self-organizes towards a truly critical state, and whether it fulfils the universal scaling relations that are conjectured for SOC<sup>12,20</sup>. In particular, the non-equilibrium critical exponents for the model described by equations (2) and (3) have not been reliably determined beyond the mean-field level, except in some limiting cases. For example, for  $b = \tau = 0$  (the number-conserving limit) the critical behaviour is governed by a critical point in the directed percolation universality class<sup>15,44</sup>. How this universality changes in the non-commuting limit of small but non-vanishing  $b$  is not conclusively understood<sup>12,20</sup>, but we expect the universality to be strongly modified, because in a renormalization group picture, the fully attractive SOC fixed point does not feature an obvious relevant direction, as is the case for directed percolation. In what follows, we experimentally implement this elusive model and provide a first experimental characterization of some of its scale-invariant properties as the system approaches the non-equilibrium critical point.

### Self-organization mechanism and model verification

We start our experiments by investigating the full time evolution of the total remaining density for different initial states. For this we prepare a gas of atoms in the ground state ( $\rho_0 = 0$ ) with different initial peak atomic densities  $n_0$  between  $0.056(5) \mu\text{m}^{-3}$  and  $0.172(2) \mu\text{m}^{-3}$ , where the numbers in parentheses refer to the standard error of the mean taken over several measurements. The Rydberg excitation laser is then suddenly switched on with  $\Omega/2\pi = 190$  kHz and  $\Delta/2\pi = 30$  MHz from the  $39p_{3/2}$  state. After an adjustable time  $t$  we turn off the excitation laser and then take an absorption image to determine  $n_t$ . Figure 2 shows that the time evolution of  $n_t$  is strikingly nonlinear and exhibits two distinct types of behaviour, depending on  $n_0$ . For high  $n_0$  there is a short initial plateau in  $n_t$ , followed by rapid exponential decay, reflecting the initial

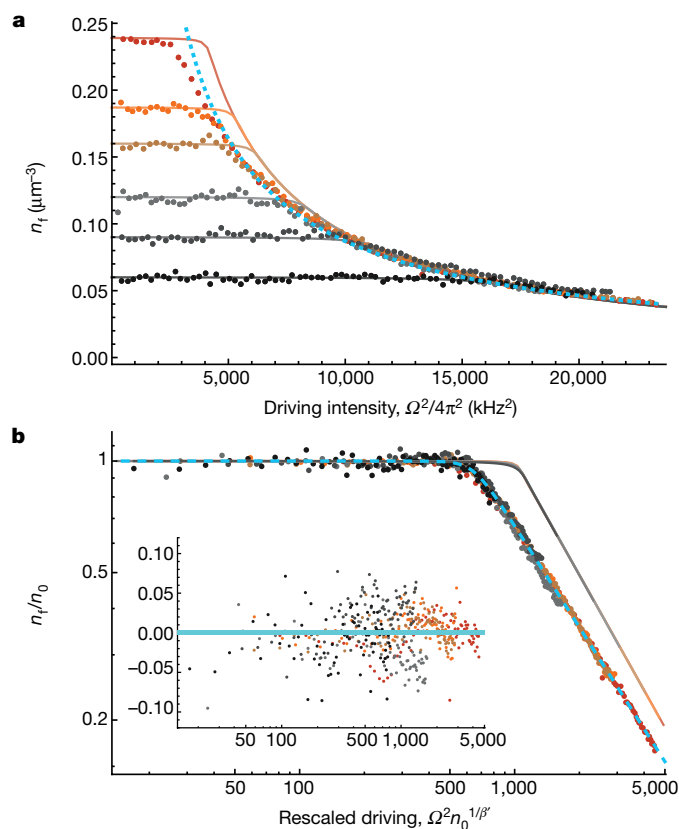
growth of the excitation density. This decay stops at a fixed density  $n_t = 0.075 \mu\text{m}^{-3}$  that is constant over a wide range of initial densities (standard deviation  $0.003 \mu\text{m}^{-3}$ ), indicating a stable attractor for the many-body dynamics. By contrast, for  $n_0 < n_t$  the dynamics appears mostly frozen, characteristic of an absorbing phase. These two types of behaviour and the sudden transition between them signal the underlying absorbing-state phase transition that depends upon the initial density and driving strength. On much longer timescales we observe a slower overall decay, which we attribute to residual single-atom excitations (and subsequent loss) with a characteristic rate  $\tau/2\pi = 1.12(2)$  Hz. Because of this slow loss, the self-organized state is not sustained indefinitely; however, the very large separation of timescales in our experiment makes it possible to robustly observe the emergent SOC features in the quasi-stationary regime (hereafter referred to as the stationary state).

We now verify that the Langevin equation provides a good theoretical description for the experimental observations. Through comparison with the data we confirm the required coupling between the active density and the total remaining density, as well as the key hierarchy of scales:  $\kappa n_0 \approx \Gamma \gg b \Gamma \gg \tau$ . For this it is sufficient to compare our data with a homogeneous mean-field approximation to the Langevin equation ( $D = 0$  and  $\xi_i = 0$ ). We find that the mean-field solutions—shown as solid lines in Fig. 2—describe the data well, except for the minor deviation in the approach to the stationary state seen around  $t \approx 2$  ms. By simultaneously fitting all of the data shown with a single set of parameters, we find  $\Gamma/2\pi = 11.7(9)$  kHz,  $\kappa/2\pi = 144(10)$  kHz  $\mu\text{m}^3$  and  $b = 0.059(5)$ , with the statistical errors estimated using bootstrap resampling. Thus the required separation of scales is satisfied by an order of magnitude or more, placing our experiments firmly in the regime in which SOC is expected. Furthermore, our experimental observations and their theoretical confirmation establish the presence of the anticipated absorbing-state phase transition and the self-organization to a stationary state that is independent of initial conditions—that is, the system displays SOC signature (1).

### Scale-invariance of the stationary density

We now turn our attention to experimental manifestations of the observed phase transition on the stationary state. In Fig. 3 we examine the dependence of the stationary density  $n_t$  (reached after  $t = 10$  ms of evolution) on the driving intensity  $\Omega^2 \propto \kappa$ . For different initial densities  $n_0$ , the stationary state exhibits a clear density-dependent critical intensity  $\Omega_c^2$  that separates the absorbing phase (with  $n_t \approx n_0$ ) from the active self-organizing phase (with  $n_t < n_0$ ). For the latter, the data fall onto a single curve resembling a power law that is independent of the initial density (dotted blue line in Fig. 3a). Although mean-field theory (solid lines) reproduces the qualitative features, the experimental data exhibit important quantitative differences, including a shift in the threshold intensity and a markedly different power-law exponent.

To further quantify the scale-invariant properties, we apply the scaling ansatz<sup>15</sup>  $n_t = n_0 F(\Omega^2 n_0^{1/\beta'})$ . By plotting  $n_t/n_0$  as a function of  $n_0^{1/\beta'} \Omega^2$ , all of the data collapse onto a single universal curve (Fig. 3b), with the best results obtained for  $\beta' = 0.869(6)$ . We find that the scaling function  $F(x)$  is well modelled by the heuristic function  $F(x) = x_c^\beta (x^{1/\beta'} + x_c^{1/\beta'})^{-1/\nu}$  (dashed blue curve in Fig. 3b), where  $x_c$  and  $\nu$  are free parameters describing the position and sharpness of the transition region between absorbing and active phases. For  $x \gg x_c$  the scaling ansatz is a power law  $n_t \propto \Omega^{-2\beta} n_0^{1-\beta/\beta'}$ , and therefore we can identify  $\beta$  as the scaling exponent that characterizes the stationary density and show that  $1-\beta/\beta'$  quantifies how (in)sensitive  $n_t$  is to the initial density. Fitting the rescaled data on a log-log scale we obtain  $\beta = 0.910(4)$ ,  $\nu = 10.6(8)$  and  $x_c = 641(3) \text{ kHz}^2 \mu\text{m}^{-3/\beta'}$ . The errors in parentheses are the standard deviation of the fitted parameters obtained via bootstrap resampling. Agreement with the scaling ansatz is confirmed by the small and evenly scattered normalized residuals between the rescaled data and the



**Fig. 3 | Scale invariance of the self-organized stationary state as a function of the driving intensity  $\Omega^2$ .** **a**, Stationary-state density  $n_i$  measured at  $t = 10$  ms as a function of  $\Omega^2$  and for different initial densities  $n_0$  using the same parameters as in Fig. 2, except with  $\Delta/2\pi = 18$  MHz. For large  $\Omega^2$  and  $n_0$ , all points collapse onto one single power-law curve  $n_i \propto \Omega^{-2\beta}$  (dotted blue line). **b**, The same data with rescaled axes to achieve full data collapse, revealing a unique scaling function (with fit shown by the dashed blue line) for the stationary density  $n_i$ . The inset shows the normalized residuals between the rescaled data and the fitted scaling function. The solid lines in **a**, **b** correspond to mean-field solutions of the effective Langevin equation. Each data point is the average of five measurements.

fitted scaling function, spanning both the absorbing and active phases (Fig. 3b, inset). The clear power-law dependence additionally rules out substantial modifications owing to the finite system size or inhomogeneous trapping geometry. Additional data taken for different densities and detunings of the driving field and slightly different experimental conditions exhibit a very similar scaling form and confirms this measured scaling exponent within an accuracy of a few per cent (Extended Data Fig. 1). By contrast, the mean-field scaling solution predicts  $\beta'_{\text{MF}} = \beta_{\text{MF}} = 1$ , which is clearly incompatible with our data. Although it is still debated to what extent SOC systems exhibit universal behaviour<sup>18,20,21</sup>, it is striking that a single function describes the stationary state over the entire accessible parameter regime, and that this function acquires a scale-invariant form characterized by a non-trivial scaling exponent—that is, SOC signature (2).

### Power-law-distributed excitation avalanches

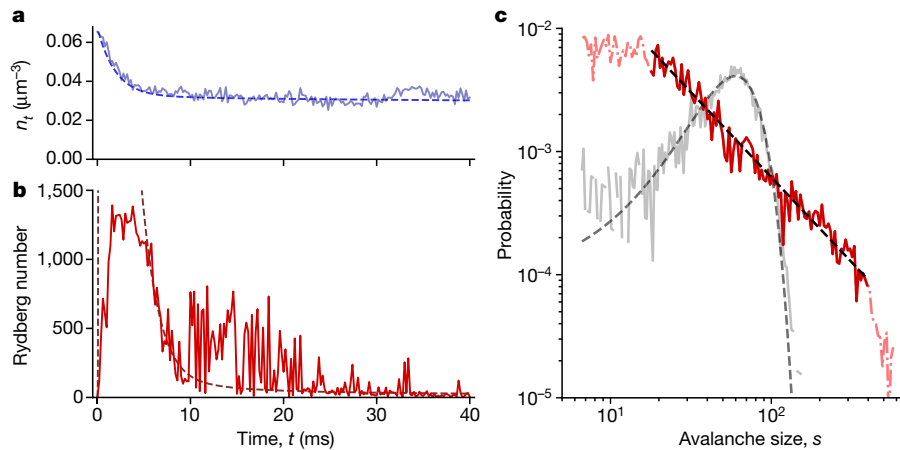
We now show that the SOC state is also evident in the statistical fluctuations of the active component. For this we use a different detection method, which is based on field ionization of the Rydberg excited atoms. For the following measurements we use the  $66p_{3/2}$  state for detection purposes, but otherwise the experimental conditions are comparable. The measurement is destructive, so each measurement point corresponds to a new experimental realization.

Figure 4a, b shows a time trace of the temporal evolution of the remaining density and the instantaneous number of excitations (the active component). The remaining density follows the same characteristic non-exponential time dependence as seen in Fig. 2, except for overall slower dynamics that can be explained by the longer lifetime and the larger  $C_6$  coefficient for the  $66p_{3/2}$  state, which lowers  $b$  in the effective description (see Supplementary Information). Figure 4b shows that the active component undergoes rapid growth at early times, which saturates the detector around 2–5 ms, and then reduces again as a consequence of the associated fast atom loss. After 10 ms the remaining density has almost reached the stationary value; however, we observe very large fluctuations of the excitation number, ranging from almost zero to clusters of up to approximately 800 excitations. We interpret this as the strong response of the system to individual excitation events that trigger avalanches that have a broad distribution of sizes and durations, which is expected as the dynamical gap vanishes close to the critical point—that is, the behaviour is evidence of SOC signature (3). In Extended Data Fig. 2 we present additional evidence of this strong response in the bulk observables following a parameter quench. The dashed lines in Fig. 4a, b show the mean-field solution to the effective Langevin equation, which describes the remaining atom number well, but as expected it completely fails to capture the large fluctuations. Additionally we observe avalanches over a wide time window (up to 40 ms) even though the remaining density appears mostly constant. This shows that the system remains close to the SOC state for an extended time period, despite the absence of an obvious particle reloading mechanism, which would be required to keep the system indefinitely at the critical point.

To investigate the distribution of the avalanche sizes  $s$ , we chose a fixed time of 25 ms and repeated the experiment 3,630 times. At this fixed time the observed excitation spikes are relatively sparse (enabling their interpretation as individual avalanche events), yet frequent enough to obtain sufficient statistics. Figure 4c shows the corresponding empirical probability-distribution function obtained by binning the data using logarithmically spaced intervals and plotted on a double logarithmic scale. The empirical probability distribution function is well described by a power law that spans 1.5 orders of magnitude and has an upper cutoff determined by the finite system size or the detector saturation (both effects are expected to play a role around  $s \approx 500$ ). The plateau around  $s \leq 20$  is attributed to the noise floor of the detector. To confirm that the observed power-law distribution is indeed a feature of the self-organizing dynamics, we also show in Fig. 4c a comparable distribution obtained by a resonant excitation pulse of 1  $\mu\text{s}$  duration, which yields a stretched Poissonian distribution, as expected for mostly uncorrelated excitations. To estimate the power-law exponent, we truncate the empirical data in the window  $20 \leq s \leq 400$  (corresponding to 2,450 measurements), and apply a maximum-likelihood estimation, yielding a power-law exponent of  $\alpha = -1.37(2)$ , where the statistical uncertainty was estimated using bootstrap resampling. The power-law exponent falls in a similar range to observations made in other conjectured SOC-like systems, such as forest fires<sup>4</sup>, neuronal networks<sup>6</sup>, earthquakes and solar flares<sup>5</sup>. However, it is important to note that non-universal corrections (owing to, for example, the non-vanishing dissipation and driving rates or imperfect separation of scales) could still affect the apparent critical properties<sup>20</sup>. Ultracold atoms offer the prospect of controlling these experimental conditions (for example, through larger detunings corresponding to lower seed excitation rates) and of determining the critical exponents for different dimensionalities in a single experimental system, permitting more stringent tests of universal scaling predictions.

The demonstrated versatility of ultracold Rydberg gases combined with the ability to understand and experimentally control the microscopic physics in this system makes it a unique platform for studying non-equilibrium collective behaviour. Future experiments could implement a mechanism to slowly add particles to the system (that is, an





**Fig. 4 | Observation of power-law distributed excitation avalanches.**

**a**, Evolution of the remaining density for the  $66p_{3/2}$  state (solid line). **b**, Simultaneously measured Rydberg atom number (active component) integrated over the whole atom cloud showing large fluctuations of the active density for  $t \geq 10$  ms (solid line). Each of the 200 plotted values corresponds to a new realization of the experiment. The dashed lines in **a**, **b** are mean-field predictions, where the effective volume of the atom cloud in **b** is adjusted for optimal agreement. **c**, Probability distribution for the instantaneous number

of Rydberg excitations (avalanche size) for 3,630 experimental runs. To determine the power-law exponent we truncate the binned data (by eye, to the region where the log–log slope is approximately constant), in a finite window indicated by the solid red line and apply a maximum-likelihood estimation. The power-law exponent  $\alpha = -1.37(2)$  is depicted by the black, straight dashed line. The grey data correspond to a control measurement for resonant excitation with a short duration laser pulse, which does not exhibit a power-law distribution. The dashed grey line is a fit to a stretched Poissonian distribution.

additional regrowth term in equation (3)) to sustain the SOC state on even longer timescales<sup>47</sup>. It should also be possible to investigate other observables beyond the mean-field level, including spatio-temporal correlations in the active and remaining densities. This would make it possible to determine multiple critical exponents and scaling relations, helping to answer long-standing questions about the universal or non-universal aspects of SOC and its relation to other non-equilibrium universality classes. Additionally, further experiments could explore the interface between driven–dissipative and isolated quantum systems governed by competing classical and quantum dynamical rules<sup>48–50</sup>, ultimately leading to a more complete and quantitative understanding of non-equilibrium universality.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-019-1908-6>.

- Bak, P., Tang, C. & Wiesenfeld, K. Self-organized criticality: an explanation of the  $1/f$  noise. *Phys. Rev. Lett.* **59**, 381–384 (1987).
- Sornette, A. & Sornette, D. Self-organized criticality and earthquakes. *EPL* **9**, 197–202 (1989).
- Rhodes, C. J. & Anderson, R. M. Power laws governing epidemics in isolated populations. *Nature* **381**, 600–602 (1996).
- Malamud, B. D., Morein, G. & Turcotte, D. L. Forest fires: an example of self-organized critical behavior. *Science* **281**, 1840–1842 (1998).
- de Arcangelis, L., Godano, C., Lippiello, E. & Nicodemi, M. Universality in solar flare and earthquake occurrence. *Phys. Rev. Lett.* **96**, 051102 (2006).
- Klaus, A., Yu, S. & Plenz, D. Statistical analyses support power law distributions found in neuronal avalanches. *PLoS ONE* **6**, e19779 (2011).
- Hesse, J. & Gross, T. Self-organized criticality as a fundamental property of neural systems. *Front. Syst. Neurosci.* **8**, 166 (2014).
- Gleeson, J. P., Ward, J. A., O’Sullivan, K. P. & Lee, W. T. Competition-induced criticality in a model of meme popularity. *Phys. Rev. Lett.* **112**, 048701 (2014).
- Shew, W. L. et al. Adaptation to sensory input tunes visual cortex to criticality. *Nat. Phys.* **11**, 659–663 (2015).
- Drossel, B. & Schwabl, F. Self-organized critical forest-fire model. *Phys. Rev. Lett.* **69**, 1629–1632 (1992).
- Dickman, R. Numerical study of a field theory for directed percolation. *Phys. Rev. E* **50**, 4404–4409 (1994).
- Muñoz, M. A., Grinstein, G., Dickman, R. & Livi, R. Critical behavior of systems with many absorbing states. *Phys. Rev. Lett.* **76**, 451–454 (1996).
- Vespignani, A. & Zapperi, S. Order parameter and scaling fields in self-organized criticality. *Phys. Rev. Lett.* **78**, 4793–4796 (1997).
- Dornic, I., Chaté, H. & Muñoz, M. A. Integration of Langevin equations with multiplicative noise and the viability of field theories for absorbing phase transitions. *Phys. Rev. Lett.* **94**, 100601 (2005).
- Henkel, M., Hinrichsen, H. & Lübeck, S. *Non-equilibrium Phase Transitions: Absorbing Phase Transitions* Vol. 1 (Springer, 2008).
- Field, S., Witt, J., Nori, F. & Ling, X. Superconducting vortex avalanches. *Phys. Rev. Lett.* **74**, 1206–1209 (1995).
- Frette, V. et al. Avalanche dynamics in a pile of rice. *Nature* **379**, 49 (1996).
- Dickman, R., Muñoz, M. A., Vespignani, A. & Zapperi, S. Paths to self-organized criticality. *Braz. J. Phys.* **30**, 27–41 (2000).
- Altshuler, E. & Johansen, T. H. Experiments in vortex avalanches. *Rev. Mod. Phys.* **76**, 471–487 (2004).
- Bonachela, J. A. & Muñoz, M. A. Self-organization without conservation: true or just apparent scale-invariance? *J. Stat. Mech.* **2009**, P09009 (2009).
- Watkins, N. W., Pruessner, G., Chapman, S. C., Crosby, N. B. & Jensen, H. J. 25 years of self-organized criticality: concepts and controversies. *Space Sci. Rev.* **198**, 3–44 (2016).
- Forster, D. *Hydrodynamic Fluctuations, Broken Symmetry, and Correlation Functions* (Taylor & Francis, 1990).
- Kosterlitz, J. M. & Thouless, D. J. Ordering, metastability and phase transitions in two-dimensional systems. *J. Phys. C* **6**, 1181 (1973).
- Frisch, U. *Turbulence: The Legacy of A. N. Kolmogorov* (Cambridge Univ. Press, 1995).
- Berges, J., Rothkopf, A. & Schmidt, J. Nonthermal fixed points: effective weak coupling for strongly correlated systems far from equilibrium. *Phys. Rev. Lett.* **101**, 041603 (2008).
- Hadzibabic, Z., Krüger, P., Cheneau, M., Battelier, B. & Dalibard, J. Berezinskii–Kosterlitz–Thouless crossover in a trapped atomic gas. *Nature* **441**, 1118–1121 (2006).
- Murthy, P. A. et al. Observation of the Berezinskii–Kosterlitz–Thouless phase transition in an ultracold Fermi gas. *Phys. Rev. Lett.* **115**, 010401 (2015).
- Tsatsos, M. C. et al. Quantum turbulence in trapped atomic Bose–Einstein condensates. *Phys. Rep.* **622**, 1–52 (2016).
- Navon, N., Gaunt, A. L., Smith, R. P. & Hadzibabic, Z. Emergence of a turbulent cascade in a quantum gas. *Nature* **539**, 72–75 (2016).
- Prüfer, M. et al. Observation of universal quantum dynamics in a spinor Bose gas far from equilibrium. *Nature* **563**, 217–220 (2018).
- Erne, S., Bücke, R., Gasenzer, T., Berges, J. & Schmiedmayer, J. Observation of universal dynamics in an isolated one-dimensional Bose gas far from equilibrium. *Nature* **563**, 225–229 (2018).
- Inouye, S. et al. Superradiant Rayleigh scattering from a Bose–Einstein condensate. *Science* **285**, 571–574 (1999).
- Clark, L. W., Gaj, A., Feng, L. & Chin, C. Collective emission of matter-wave jets from driven Bose–Einstein condensates. *Nature* **551**, 356 (2017).
- Ho, T.-L. Universal thermodynamics of degenerate quantum gases in the unitarity limit. *Phys. Rev. Lett.* **92**, 090402 (2004).
- Lesanovsky, I. & Garrahan, J. P. Kinetic constraints, hierarchical relaxation, and onset of glassiness in strongly interacting and dissipative Rydberg gases. *Phys. Rev. Lett.* **111**, 215305 (2013).
- Carr, C., Ritter, R., Wade, C. G., Adams, C. S. & Weatherill, K. J. Nonequilibrium phase transition in a dilute Rydberg ensemble. *Phys. Rev. Lett.* **111**, 113901 (2013).
- Schempp, H. et al. Full counting statistics of laser excited Rydberg aggregates in a one-dimensional geometry. *Phys. Rev. Lett.* **112**, 013002 (2014).

38. Malossi, N. et al. Full counting statistics and phase diagram of a dissipative Rydberg gas. *Phys. Rev. Lett.* **113**, 023006 (2014).
39. Urvoy, A. et al. Strongly correlated growth of Rydberg aggregates in a vapor cell. *Phys. Rev. Lett.* **114**, 203002 (2015).
40. Valado, M. M. et al. Experimental observation of controllable kinetic constraints in a cold atomic gas. *Phys. Rev. A* **93**, 040701 (2016).
41. Goldschmidt, E. A. et al. Anomalous broadening in driven dissipative Rydberg systems. *Phys. Rev. Lett.* **116**, 113001 (2016).
42. Simonelli, C. et al. Seeded excitation avalanches in off-resonantly driven Rydberg gases. *J. Phys. B* **49**, 154002 (2016).
43. Letscher, F., Thomas, O., Niederprüm, T., Fleischhauer, M. & Ott, H. Bistability versus metastability in driven dissipative Rydberg gases. *Phys. Rev. X* **7**, 021020 (2017).
44. Gutiérrez, R. et al. Experimental signatures of an absorbing-state phase transition in an open driven many-body quantum system. *Phys. Rev. A* **96**, 041602 (2017).
45. Ates, C., Pohl, T., Pattard, T. & Rost, J. M. Many-body theory of excitation dynamics in an ultracold Rydberg gas. *Phys. Rev. A* **76**, 013413 (2007).
46. Marcuzzi, M., Schick, J., Olmos, B. & Lesanovsky, I. Effective dynamics of strongly dissipative Rydberg gases. *J. Phys. A* **47**, 482001 (2014).
47. Klocke, K. & Buchhold, M. Controlling excitation avalanches in driven Rydberg gases. *Phys. Rev. A* **99**, 053616 (2019).
48. Marcuzzi, M., Buchhold, M., Diehl, S. & Lesanovsky, I. Absorbing state phase transition with competing quantum and classical fluctuations. *Phys. Rev. Lett.* **116**, 245701 (2016).
49. Buchhold, M., Everest, B., Marcuzzi, M., Lesanovsky, I. & Diehl, S. Nonequilibrium effective field theory for absorbing state phase transitions in driven open quantum spin systems. *Phys. Rev. B* **95**, 014308 (2017).
50. Pérez-Espigares, C., Marcuzzi, M., Gutiérrez, R. & Lesanovsky, I. Epidemic dynamics in open quantum spin systems. *Phys. Rev. Lett.* **119**, 140401 (2017).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2020



## Methods

### Sample preparation

Our experiments are performed using a thermal gas of potassium-39 atoms, loaded directly from a magneto-optical trap into a crossed optical dipole trap. The resulting cigar shaped atom cloud has a temperature of 40  $\mu$ K and  $e^{-1/2}$  radii of  $10 \mu\text{m} \times 100 \mu\text{m}$ . This should be compared to the characteristic distance between facilitated Rydberg excitations  $r_{\text{fac}} = (C_6/\Delta)^{1/6}$ , which for a detuning of  $\Delta/2\pi = 30$  MHz is about  $1.7 \mu\text{m}$ . The peak number of atoms in the  $|g\rangle$  state is  $1.3 \times 10^5$ , and the density determined by in situ imaging is  $2.4 \times 10^{11} \text{ cm}^{-3}$ . To vary the density while holding all other parameters fixed, we reduce the magneto-optical trap loading time. The lifetime of the atoms in the trap without Rydberg excitation is about 4 s—that is, much longer than the relevant timescales for the SOC dynamics.

### Excitation laser

To excite the atoms to the  $39p_{3/2}$  Rydberg state we use a single photon optical transition at a wavelength of 285 nm. This light is produced by a frequency-doubled dye laser delivering up to 200 mW of single-mode light and is frequency stabilized to a high-finesse cavity, resulting in an independently measured linewidth of 400 kHz. The excitation beam is aligned parallel to the long axis of the trap and weakly focused to a waist much larger than the size of the atom cloud such that it is practically uniform. We experimentally determine the Rabi frequency  $\Omega$  for every individual repetition of the experiment by logging the respective single-shot-laser power on a photodiode and employing an independent Rabi frequency calibration based on measuring the light shifts induced by the laser via Ramsey interferometry<sup>51</sup>.

### Numerical simulation of the Langevin equation

Although the Langevin equation (equations (2) and (3)) is straightforward to solve in the mean-field approximation, in Fig. 1 we show exemplary numerical simulations that capture the effects of diffusion and multiplicative noise terms in a one-dimensional setting. For these simulations we make use of the XMDS2 (stochastic) differential-equation solver package<sup>52</sup>, assuming a transverse grid size of 128 points and a timestep of  $2.5 \times 10^{-3}$ . The noise term is implemented as a zero-mean Wiener process with a standard deviation proportional to  $\sqrt{p}$ . However, to ensure numerical stability we found it necessary to impose a noise cutoff by setting  $\xi = 0$  when  $p < 0.0025n_0$ . For  $b = 0$  the solutions exhibit an absorbing-state phase transition at  $n_0 = 0.39$  and power-law scaling consistent with directed percolation universality (in one-dimension,  $\beta_{\text{DP}} = 0.276$ ). For  $b \neq 0$  we find that the individual timetraces obtained from the full numerical solution are qualitatively very similar to the corresponding mean-field solutions. By fitting the numerical results in the same manner as performed for the experimental data, we obtain slightly larger effective parameters  $\kappa$  and  $\Gamma$ .

### Comparison of the power-law hypothesis to alternative distributions

To test whether the avalanche data are indeed described by a power-law distribution we employ the widely used Kolmogorov–Smirnov (KS) test against several alternative distributions, including other heavy-tailed distributions (following the definitions in ref. 6). The KS statistic is defined as the maximum distance between the cumulative distribution of the empirical data and that of the hypothesized distribution, with small values much less than 1 indicating good agreement. In all cases we minimize the KS statistic as a function of the parameters of the hypothesized distribution, restricting the data and the hypothesized distributions to the range  $20 \leq s \leq 400$ . For the data depicted in Fig. 4, the obtained KS-test statistics are: 0.015 (power law), 0.102 (exponential), 0.031 (log-normal), and 0.04 (gamma). This shows that the power-law distribution provides a better fit to the data than the alternative distributions. The power-law exponent  $\alpha = -1.38$  found via

KS minimization is in excellent agreement with the value obtained via the maximum-likelihood estimation<sup>53</sup>.

### Detuning dependence and further evidence for non-equilibrium universality

In the following we present additional evidence for the universal nature of the self-organized stationary state. For this we performed additional measurements of the stationary density as a function of the driving intensity but for different detunings of the excitation laser, as shown in Extended Data Fig. 1. Each dataset shows qualitatively similar behaviour to that presented in Fig. 3, clearly showing the transition from an absorbing phase to a self-organizing active phase. However, these data also show that the location of the critical point depends on the laser detuning.

To further analyse these data we apply the scaling ansatz  $n_f = n_0 F(\Omega^2 \Delta^d n_0^{1/\beta'})$ , where  $\beta' = 0.869$ , and we have included as a new parameter the detuning rescaling exponent,  $d$ . For  $d = -2.06(1)$  the data again collapse onto a single universal curve. In this way we determine the  $\kappa \propto \Omega^2/\Delta^{2.06}$  dependence of the spreading parameter, used elsewhere in the paper to compare the data with mean-field theory.

Before analysing the scaling properties of the rescaled data, careful inspection shows that it has a slightly different form to the scaling function  $F(x)$  used to describe the data in Fig. 3b. This is evidenced by the fit to  $F(x)$ , shown as a blue dashed line in Extended Data Fig. 1b. The deviation is most apparent in the normalized fit residuals (Extended Data Fig. 1, inset) which, in contrast to Fig. 3b, exhibits some structure (for example, the inverted U-shape of the black datapoints). Unless properly accounted for, this deviation between the scaling form of the data and the heuristic scaling function causes a systematic error in the determination of the critical scaling exponent. To rectify this, we model the detuning-dependent data by a generalized scaling function  $F'(x) = [1 + (x/x_c)^{\alpha} + (x/x_c)^{\beta}]^{-1/\nu}$ , where the newly introduced parameters  $x_c < x_c$  and  $\alpha < \beta$  empirically describe power-law scaling for intermediate driving intensities. In the asymptotic regime  $x \gg x_c$ , the scaling function once again reduces to a power law  $n_f/n_0 \propto x^{-\beta}$ .

### Critical response

As additional evidence for the system reaching a critical state, we have investigated the gapless response of the stationary state following a parameter quench. Assuming the SOC state is indeed an attractor for the dynamics, on one hand we expect that small perturbations (for example a sudden change of the spreading parameter  $\kappa$ ), should trigger avalanche-like processes that eventually bring the system back to a new critical state corresponding to a lower stationary density. On the other hand, if the system evolves to a state that is deep within the absorbing phase, then avalanches can only be triggered by perturbations larger than a threshold value corresponding to a non-zero dynamical gap. To measure this response we start from the stationary state (reached after  $t = 10$  ms) corresponding to different driving intensities  $\Omega_i^2$  (sketched in Extended Data Fig. 2a). We then perturb the system by quenching the driving intensity to a new value  $\Omega_f^2$  and then wait for a further 10 ms before measuring the new stationary density. The whole procedure is then repeated with a slightly larger final driving intensity  $\Omega_f^2 \approx \Omega_i^2 + (2\pi \times 50 \text{ kHz})^2$ . From these two measurements we estimate the susceptibility  $\chi = dn_f/d\Omega_f^2 = [n_f(\Omega_f^2) - n_f(\Omega_i^2)]/(\Omega_f^2 - \Omega_i^2)$ .

Extended Data Fig. 2 shows the measured susceptibility as a function of  $\delta = (\Omega_f^2 - \Omega_i^2)/\Omega_c^2$  for three different initial conditions corresponding to  $\Omega_i < \Omega_c$  (absorbing),  $\Omega_i \approx \Omega_c$  (critical) and  $\Omega_i > \Omega_c$  (active). For each of these initial conditions we observe pronounced minima in  $\chi$  corresponding to the strongest system response. We interpret the leading edge on the left side of each minimum as the point where the perturbation is sufficient to bring the system back to the active phase, thereby triggering avalanche-like dynamics and extra loss. When starting deep in the absorbing phase (black circles) the onset occurs at a large value of  $\delta$ , which is a measure of the non-zero dynamical gap. By contrast,

the onsets for critical (brown triangles) and active (red squares) initial states both coincide at  $\delta \approx 0$  within the experimental resolution. We can compare these data to a prediction of the susceptibility obtained from a derivative of the experimentally determined scaling function using  $\beta = 0.910$ . The scaling-function predictions (solid lines in Extended Data Fig. 2) are in good agreement with the data, whereas mean-field predictions (dotted lines) systematically fail to capture the widths and heights of the peaks. However, we note that starting from the initial active state, the measured response is narrower and slightly stronger than the scaling-function prediction. This is further evidence that the system evolves towards a state that is sharply concentrated at the critical point, instead of one that is a statistical mixture of many different accessible macro states. From these experiments we confirm that when starting from a supercritical state (irrespective of the precise value of  $\Omega_i > \Omega_c$ ), the system self-organizes to a critical state that is characterized by a vanishing excitation gap—underpinning SOC signature (3).

Fitting the generalized scaling function to the rescaled data yields  $\beta = 0.95(3)$ , where the larger statistical uncertainty reflects the fact that the generalized model function has more parameters. This is close to the value  $\beta = 0.910(4)$  determined from the density-dependent data in the main text. Refitting the density-dependent data with the generalized scaling function yields  $\beta = 0.920(7)$ . This shows that although the full form of the scaling function is not universal, data taken under very different conditions concerning initial densities and detuning of the driving field do in fact share a common universal critical exponent describing the asymptotic scaling regime.

## Role of trap inhomogeneities and residual coherence

We can also rule out possible modifications to the scaling behaviour due to other experimental details such as the inhomogeneous density or residual effects of quantum coherence.

**Inhomogeneities.** In the experiment the atoms are laser-trapped in a cylindrical geometry of finite diameter and length, causing a nearly homogeneous density distribution in the trap centre and smooth variation of  $n_i$  at the boundaries.  $n_i$  smoothly follows the Gaussian trapping profile of the lasers. To estimate the impact of inhomogeneities, on this basis we now study a local density approximation for the Langevin equation. In this approximation,  $\rho(\mathbf{r}, t)$  experiences a constant background density  $n(\mathbf{r}, t) = \tilde{n}(\mathbf{r}, t)I(\mathbf{r})$  at each point in space  $\mathbf{r}$ , which is modulated by the trapping profile  $I(\mathbf{r})$ , whereas  $\tilde{n}(\mathbf{r}, t)$  only incorporates fluctuations owing to the coupling to  $\rho_i$ . An appropriate mean-field theory considers  $\rho_i = V^{-1} \int_V \rho(\mathbf{r}, t)$  as the spatially averaged density over the system volume  $V$ , and  $\tilde{n}_{t=0} = n_0$  in the absence of fluctuations. The corresponding spatially averaged SOC line is located at  $\Omega_c^2 \propto \kappa_c = (I/n_0) \int_V [1/I(\mathbf{r})] \propto n_0^{-1}$ , demonstrating that the mean-field exponent  $\beta = 1$  is not modified by the inhomogeneous geometry.

**Quantum coherence.** The evolution of the density averages (Supplementary Equations (S4) and (S5)) is real and linear in time,

which maps the final Langevin equation to a stochastic differential equation for classical processes. It incorporates strong, classical correlations between different atoms but lacks the possibility for long range coherence. Coherence between different atoms can be systematically built-in by replacing the adiabatic elimination (Supplementary Equation (S1)) with the exact solution, which amounts to a shift  $\Gamma \rightarrow \Gamma + \partial_t$  (Supplementary Equation (S5)). To leading order it introduces a coherent contribution,  $(\Gamma + \gamma_{dc})^{-1} \partial_t^2 m_l$ , to the right-hand side of Supplementary Equation (S5), where  $m_l$  is the probability for having an excited particle within the coarse grained region  $l$  corresponding to the characteristic facilitation volume. Analogously to a damped harmonic oscillator, this evolution is observable on timescales  $t(\Gamma + \gamma_{dc}) \leq 1$ , but washed out on larger timescales, that is, on the relaxation towards the SOC steady state. Fast coherent processes might modify the parameters  $\kappa$ ,  $D$  and  $\tau$ , but not the structure of the Langevin equation.

## Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

51. Arias, A., Lochead, G., Wintermantel, T. M., Helmrich, S. & Whitlock, S. Realization of a Rydberg-dressed Ramsey interferometer and electrometer. *Phys. Rev. Lett.* **122**, 053601 (2019).
52. Dennis, G. R., Hope, J. J. & Johnsson, M. T. XMD2: fast, scalable simulation of coupled stochastic partial differential equations. *Comput. Phys. Commun.* **184**, 201–208 (2013).
53. Bauke, H. Parameter estimation for power-law distributions by maximum likelihood methods. *Eur. Phys. J. B* **58**, 167–173 (2007).

**Acknowledgements** We acknowledge T. Ebbesen, G. Pupillo and M. Weidemüller for discussions. This work is supported by the Deutsche Forschungsgemeinschaft under WH141/1-1 and is part of and supported by the DFG Collaborative Research Centre ‘SFB 1225 (ISOQUANT)’, the Heidelberg Center for Quantum Dynamics, the European Union H2020 FET Proactive project RySQ (grant number 640378) and the ‘Investissements d’Avenir’ programme through the Excellence Initiative of the University of Strasbourg (IdEx). M.B. acknowledges support from the Alexander von Humboldt Foundation. S.D. acknowledges support by the German Research Foundation (DFG) through the Institutional Strategy of the University of Cologne within the German Excellence Initiative (ZUK 81) and the European Research Council via ERC grant agreement number 647434 (DOQS). S.W. was partially supported by the University of Strasbourg Institute for Advanced Study (USIAS). S.H. acknowledges support by the Carl Zeiss Foundation, A.A. and S.H. acknowledge support by the Heidelberg Graduate School for Fundamental Physics.

**Author contributions** S.H., G.L. and S.W. devised the experiments; S.H., A.A., G.L. and T.M.W. acquired and analysed the data; M.B. and S.D. developed the theoretical description; all authors contributed to interpreting the results and writing of the manuscript.

**Competing interests** The authors declare no competing interests.

## Additional information

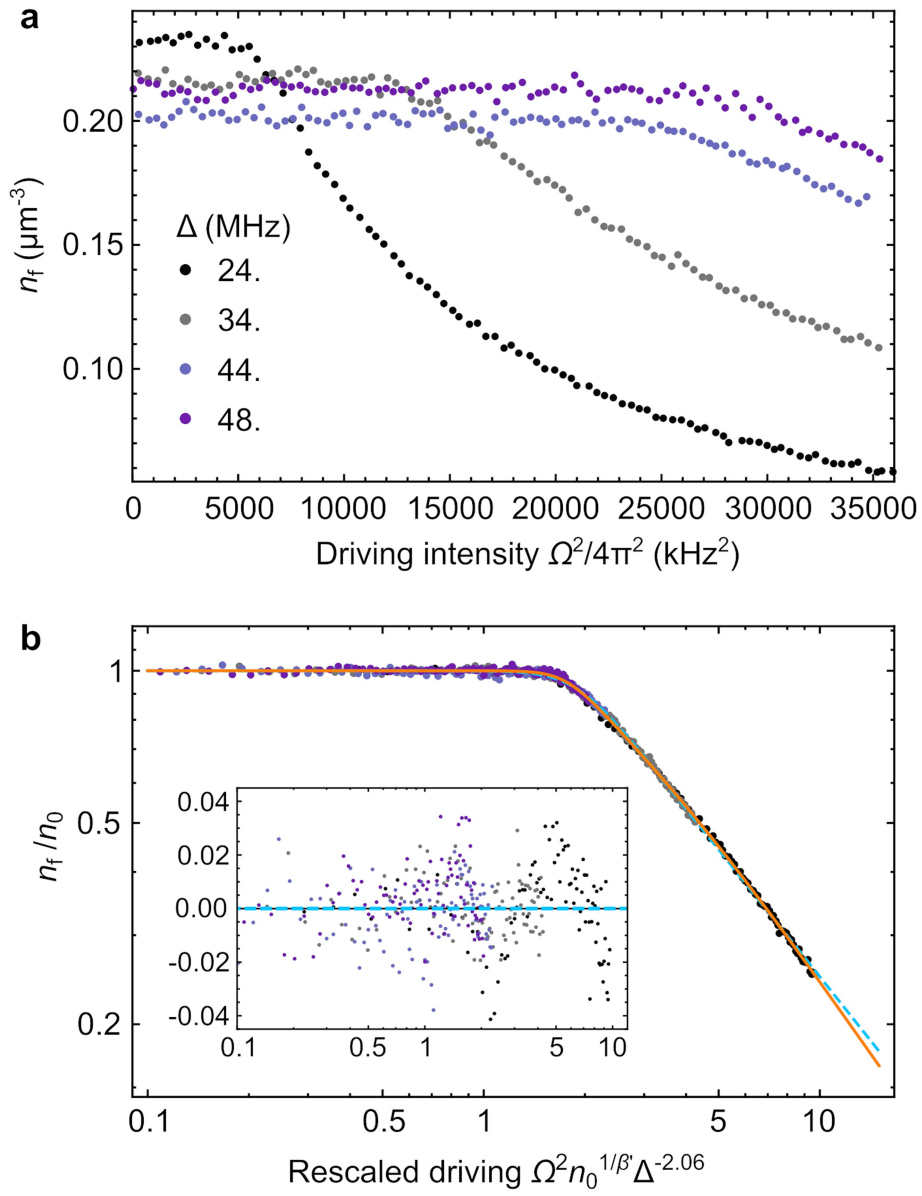
**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41586-019-1908-6>.

**Correspondence and requests for materials** should be addressed to S.W.

**Peer review information** *Nature* thanks Ronald Dickman and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

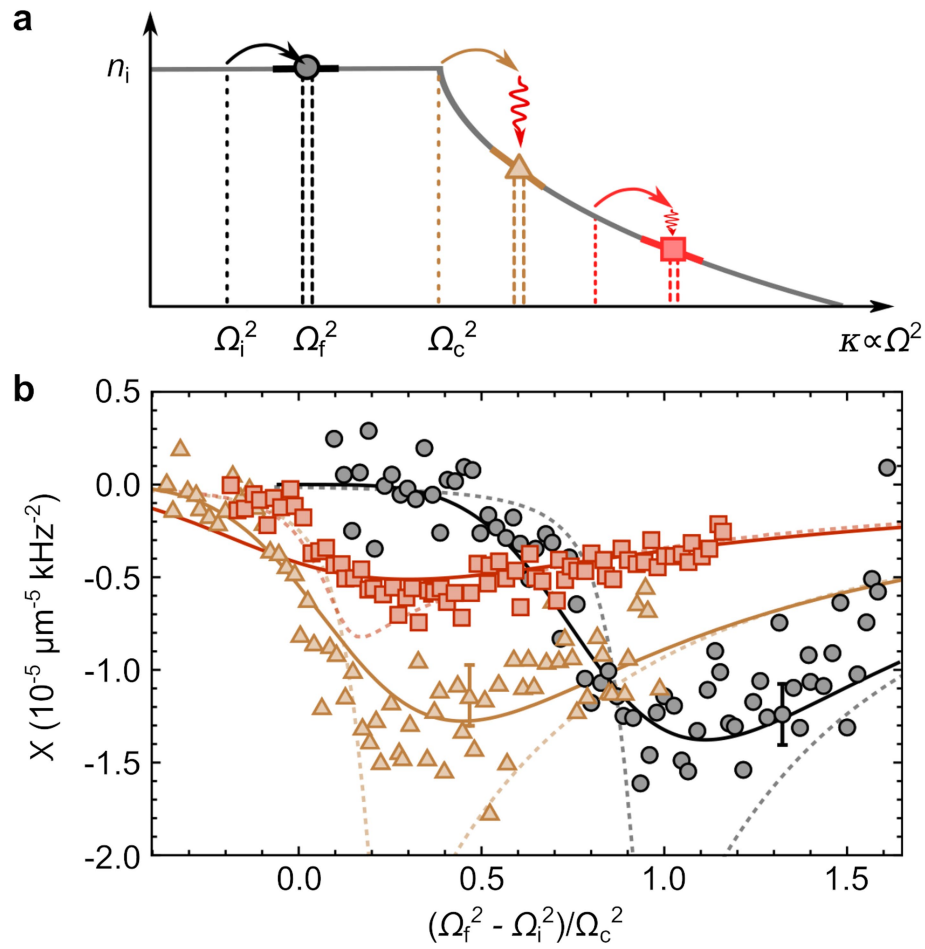
**Reprints and permissions information** is available at <http://www.nature.com/reprints>.





**Extended Data Fig. 1 | Further evidence for non-equilibrium universality.**  
**a**, Stationary-state density  $n_f$  measured at  $t = 10$  ms as a function of  $\Omega^2$  and for different detunings  $\Delta$ . **b**, The same data with rescaled axes to achieve full data collapse, revealing the scaling function (with fit shown by the dashed blue line) for the stationary density  $n_f$ . Inset, normalized residuals between the rescaled

data and the fitted scaling function. The dashed blue line corresponds to the simple scaling function used in the main text, and the solid orange line is a generalized scaling function that reproduces the asymptotic scaling form more accurately. Each data point corresponds to a single measurement.



**Extended Data Fig. 2 | Response of the SOC state to external perturbations.**

**a.** Sketch of the experimental procedure used to measure the susceptibility  $\chi = dn_i/d\Omega_i^2$  by quenching the spreading parameter  $\kappa \propto \Omega^2$  across the absorbing state phase transition. **b.** Experimental data corresponding to three different initial conditions corresponding to the absorbing phase (black circles), critical phase (brown triangles) and active phase (red squares). The solid lines

correspond to predictions based on the experimentally determined scaling function, and the dotted lines correspond to mean-field predictions. Each data point corresponds to the average of eight measurements. For reference we show two representative error bars, corresponding to the standard error of the mean.



# Fast two-qubit logic with holes in germanium

<https://doi.org/10.1038/s41586-019-1919-3>

N. W. Hendrickx<sup>1,2,4</sup>, D. P. Franke<sup>1,2,4</sup>, A. Sammak<sup>1,3</sup>, G. Scappucci<sup>1,2</sup> & M. Veldhorst<sup>1,2\*</sup>

Received: 17 June 2019

Accepted: 8 October 2019

Published online: 13 January 2020

Universal quantum information processing requires the execution of single-qubit and two-qubit logic. Across all qubit realizations<sup>1</sup>, spin qubits in quantum dots have great promise to become the central building block for quantum computation<sup>2</sup>. Excellent quantum dot control can be achieved in gallium arsenide<sup>3–5</sup>, and high-fidelity qubit rotations and two-qubit logic have been demonstrated in silicon<sup>6–9</sup>, but universal quantum logic implemented with local control has yet to be demonstrated. Here we make this step by combining all of these desirable aspects using hole quantum dots in germanium. Good control over tunnel coupling and detuning is obtained by exploiting quantum wells with very low disorder, enabling operation at the charge symmetry point for increased qubit performance. Spin–orbit coupling obviates the need for microscopic elements close to each qubit and enables rapid qubit control with driving frequencies exceeding 100 MHz. We demonstrate a fast universal quantum gate set composed of single-qubit gates with a fidelity of 99.3 per cent and a gate time of 20 nanoseconds, and two-qubit logic operations executed within 75 nanoseconds. Planar germanium has thus matured within a year from a material that can host quantum dots to a platform enabling two-qubit logic, positioning itself as an excellent material for use in quantum information applications.

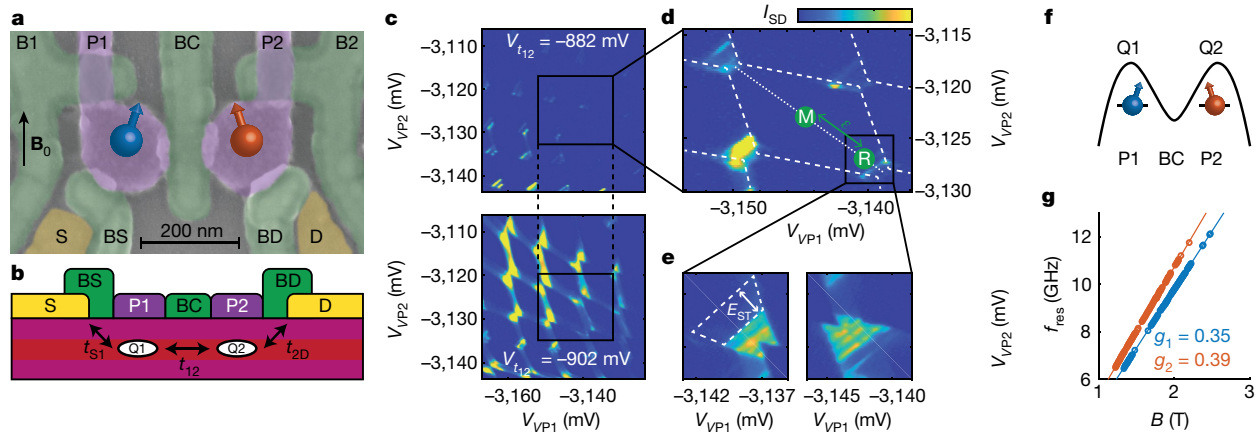
Gate-defined quantum dots were recognized early on as a promising platform for quantum information<sup>2</sup> and many materials have been investigated as hosts for the quantum dots. Initial research mainly focused on the low-disorder semiconductor gallium arsenide<sup>10,11</sup>. Steady progress in the control and understanding of this system culminated in the initial demonstration and optimization of spin qubit operations<sup>12</sup> and the realization of rudimentary analogue quantum simulations<sup>3</sup>. However, the omnipresent hyperfine interactions in group III–V materials seriously deteriorate the spin coherence. Considerable improvements to the coherence times could be achieved by switching to the group IV semiconductor silicon, in particular when defining spin qubits in an isotopically purified host crystal with vanishing concentrations of non-zero nuclear spins<sup>13</sup>. This enabled single-qubit rotations with fidelities beyond 99.9%<sup>7</sup> and the execution of two-qubit logic gates with fidelities up to 98%<sup>6,8,9</sup>, underlining the potential of spin qubits for quantum computation. Nevertheless, quantum dots in silicon are often formed at unintended locations, and control over the tunnel coupling determining the strength of two-qubit interactions is limited. Moreover, the absence of a sizable spin–orbit coupling for electrons in silicon requires the inclusion of microscopic components such as on-chip striplines or nanomagnets close to each qubit, which complicates the design of large and dense two-dimensional (2D) structures. Scalability thus remains a challenge for these systems, and a platform that can overcome these limitations would be highly desirable.

Hole states in semiconductors<sup>14</sup> typically exhibit strong spin–orbit coupling (SOC), which has enabled the demonstration of fast single-qubit rotations<sup>15–17</sup>. Furthermore, whereas valley degeneracy

complicates qubit definition for electrons in silicon, this is absent for holes, and excited states can be well separated in energy. In silicon, unfavourable band alignment prevents strain engineering of low-disorder quantum wells for holes, restricting experiments to metal–oxide–semiconductor structures<sup>18</sup>. Research on germanium has mostly focused on self-assembled nanowires<sup>19</sup> and has demonstrated single-shot spin readout<sup>20</sup> and coherent spin control<sup>17</sup>. However, strained germanium can reach hole mobilities<sup>21</sup> of  $\mu > 10^6 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$ , and undoped germanium quantum wells were recently shown to support the formation of gate-controlled hole quantum dots<sup>22</sup>. Now, the crucial challenge is the demonstration of coherent control in this platform and the implementation of qubit–qubit gates for scalable quantum information with holes.

Here we make this step and demonstrate single- and two-qubit logic with holes in planar germanium. We fabricate devices on silicon substrates, using standard manufacturing materials. We grow undoped strained germanium quantum wells, measured to have high hole mobilities  $\mu > 5 \times 10^5 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$  and a low effective hole mass<sup>22,23</sup>  $m_h = 0.09 m_e$ , extrapolated to reach  $m_h = 0.05 m_e$  at zero density<sup>24</sup>, with  $m_e$  the electron rest mass. This allows us to define quantum dots of comparatively large size, and we find excellent control over the exchange interaction between the two dots. We operate in a multi-hole mode, reducing challenges in tuning and characterization, which is advantageous for scaling. We make use of the spin–orbit interaction for qubit driving and perform single-qubit rotations at frequencies exceeding 100 MHz. This advantage of fast driving becomes further apparent in coherently accessing the Hilbert space of a two-qubit system.

<sup>1</sup>QuTech, Delft University of Technology, Delft, The Netherlands. <sup>2</sup>Kavli Institute of Nanoscience, Delft University of Technology, Delft, The Netherlands. <sup>3</sup>Netherlands Organisation for Applied Scientific Research (TNO), Delft, The Netherlands. <sup>4</sup>These authors contributed equally: N. W. Hendrickx, D. P. Franke. \*e-mail: m.veldhorst@tudelft.nl



**Fig. 1 | Fabrication and operation of a planar germanium double quantum dot.** **a**, False-coloured scanning electron microscope image of the two-qubit device, where Ohmic contacts are indicated in yellow, the barrier gate layer is depicted in green and the plunger gate layer in purple. Two hole quantum dots, indicated by the blue and red arrows, are formed in a high-mobility Ge quantum well and controlled by the electric gates. The direction of the external field  $B_0$  is indicated by the black arrow. **b**, Schematic cross-section of the system, where quantum dots are formed below plunger gates P1 and P2, while the different tunnelling rates can be controlled by barrier gates BS, BD and BC. **c**, Transport current through the double dot as a function of plunger gate voltages for weak (top) and strong (bottom) interdot coupling, mediated by a virtual tunnel gate. **d**, Charge stability diagram of the qubit operation point, where the dashed lines

correspond to the charge transitions. The detuning axis  $\epsilon$  is indicated by the dotted line, with label R corresponding to the qubit readout point. To allow coherent control of the isolated spin states, a two-level voltage pulse on gates P1 and P2 is used to detune the dot potentials and prevent tunnelling to and from the dots during the manipulation phase (label M). **e**, Transport current through the double dot as a function of plunger gate voltage for positive (left) and negative (right) bias. Pauli spin blockade becomes apparent from the suppression of the transport current for the positive bias direction, up to the singlet–triplet energy splitting of  $E_{ST} = 0.6$  meV. **f**, Illustration of the energy landscape in our double-quantum dot system. **g**, Resonance frequency,  $f_{\text{res}}$ , of the two qubits as a function of the external magnetic field, showing the individual qubit resonances.

For example, in silicon the execution of a controlled NOT (CNOT) gate implemented with an on-chip stripline has been shown using microsecond long pulses<sup>6,8</sup>, and this timescale can be reduced to 0.2–0.5  $\mu\text{s}$  by incorporating nanomagnets<sup>9</sup>. Here we demonstrate that the spin–orbit coupling of holes in germanium together with the sizable exchange interaction enables a CNOT within 75 ns.

A scanning electron microscope image of the germanium two-qubit device is shown in Fig. 1a. To accumulate holes and define two quantum dots, the circular plunger gates are set to negative potential ( $V_{P1}$ ,  $V_{P2} \approx -2$  V). The tunnel coupling between the dots  $t_{12}$  and the tunnel couplings to the source and drain reservoirs ( $t_{1S}$ ,  $t_{2D}$ ) are controlled by the barrier gates BC, BS and BD, respectively. Working in a virtual gate voltage space ( $V_{VP1}$ ,  $V_{VP2}$ ,  $V_{t12}$ ,  $V_{t2D}$  and  $V_{t1S}$ ), we can independently tune these properties (see Supplementary Videos 1–3 online for video-mode operation). We measure the transport current through the double dot system (Fig. 1c, d), and for certain hole occupations (Extended Data Fig. 3) we observe a suppression of the transport current for a positive bias voltage  $V_{SD} = 1$  mV, caused by Pauli spin blockade (PSB) (see Fig. 1e). We make use of the blockade as an effective method for spin-to-charge conversion<sup>2,11</sup>, as well as to initialize our two-qubit system in the blocked  $|\downarrow\downarrow\rangle$  ground state.

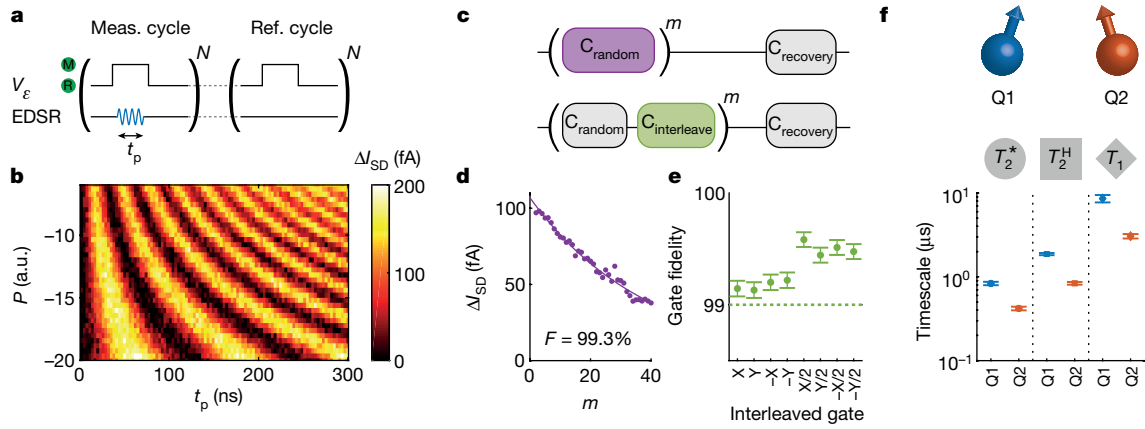
Taking advantage of the strong spin–orbit coupling<sup>17</sup>, we are able to implement a fast manipulation of the qubit states by electric dipole spin resonance (EDSR). We tune the device to a readout point within the PSB region (indicated by the label R in Fig. 1d) and apply an electric microwave excitation to gate P1. When the frequency of the microwave excitation matches the spin resonance frequency of either qubit, PSB is lifted and an increase in the transport current can be observed. We extract the resonance frequency of each qubit as a function of external magnetic field strength  $B_0$  (Extended Data Fig. 4) and observe two distinct qubit resonance lines with  $g$ -factors  $g_1 = 0.35$  and  $g_2 = 0.38$  (Fig. 1g). The difference in  $g$ -factors between the two dots is likely to be caused by slightly different hole fillings and thus quantum dot orbitals. As an effect of the spin–orbit coupling, a strong orbital dependence of the effective  $g$ -factor is typically measured in hole quantum dots<sup>18,25</sup>. Furthermore, the effective  $g$ -factor can be tuned electrically as a direct

result of the SOC<sup>26</sup> (see for example Fig. 3c, d), thereby guaranteeing independent control of the different qubits. We observe that the resonance frequency of both qubits remains stable over several hours, with discrete jumps at longer timescales as presented in Extended Data Fig. 5.

We developed a measurement technique in which we measure the averaged transport current over  $N$  repeated pulse cycles and subtract a reference measurement using a lock-in amplifier, to mitigate slow variations in the transport current (see Methods), as is indicated in Fig. 2a. After readout, the system is left in the blocking  $|\downarrow\downarrow\rangle$  state, serving as the initialization of our qubits. We now operate the device in the single-qubit transport mode in an external field of  $B_0 = 0.5$  T and use the second qubit (Q2) as a readout ancilla. Coherent control over the qubit is demonstrated in a Rabi experiment, where the spin state of qubit 1 (Q1) is measured as a function of microwave pulse length  $t_p$  and power  $P$ , as shown in Fig. 2b. By increasing the power of the microwave pulse, we can reach Rabi frequencies of over 100 MHz, at an elevated field  $B_0 = 1.65$  T (Extended Data Fig. 6).

To determine the control fidelity, which describes the accuracy of our quantum gates, we implement randomized benchmarking of the single-qubit Clifford group<sup>27</sup> (Fig. 2c). The measured decay curve of the qubit state as a function of sequence length  $m$  is shown in Fig. 2d, from which we extract a single-qubit control fidelity of  $F_c = 99.3\%$ , using gate times  $t_\pi = 20$  ns and  $t_{\pi/2} = 10$  ns. In Fig. 2e, we show the gate fidelities for the different  $\pi$  and  $\pi/2$  gates as obtained by interleaved randomized benchmarking, where each randomly drawn gate is followed by the respective interleaved gate (see Fig. 2c). All individual gate fidelities are  $F_c > 99\%$ , with the infidelity for  $\pi/2$  gates being approximately twice as low as for the  $\pi$  gates, on account of the difference in pulse length.

We extensively characterize the coherence in our system at an exchange coupling of  $J/h \approx 20$  MHz and find  $T_{2,Q1}^* = 833$  ns and  $T_{2,Q2}^* = 419$  ns, which can be extended by performing a Hahn echo to  $T_{2,Q1}^H = 1.9$   $\mu\text{s}$  and  $T_{2,Q2}^H = 0.8$   $\mu\text{s}$  (data in Extended Data Fig. 7), as indicated in Fig. 2f. These coherence times compare favourably to  $T_2^* = 130$  ns for germanium hut wires<sup>17</sup> and  $T_2^* = 270$  ns for holes in silicon<sup>28</sup>. Electrons in GaAs have an even shorter dephasing time<sup>11</sup>, with  $T_2^* = 10$  ns. The limited  $T_2^*$  in GaAs is due to hyperfine interactions,



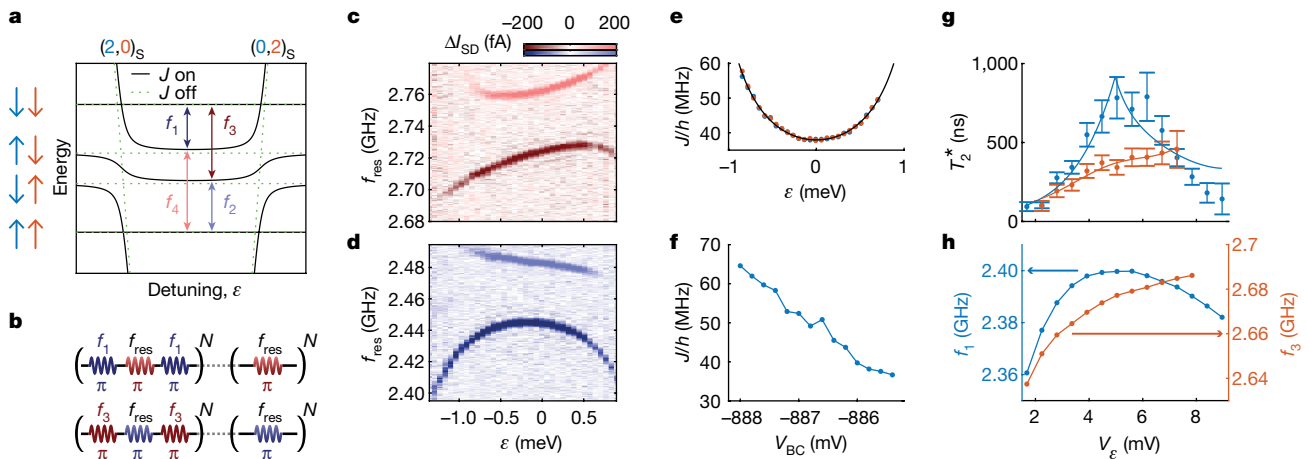
**Fig. 2 | Qubit control, gate fidelity and quantum coherence of planar germanium qubits.** **a**, Measurement sequence used for the Rabi driving measurements. Measurement cycles with EDSR pulses are alternated with reference cycles without a microwave tone, allowing an efficient background current subtraction. Each cycle is repeated  $N$  times, such that measurement and reference cycles alternate at a typical lock-in frequency of  $f_{\text{meas}} = 89.75$  Hz. **b**, Colour map of the differential bias current  $\Delta I_{\text{SD}}$  as a function of microwave pulse time  $t_p$  and power  $P$ , where clear Rabi rotations on Q1 can be

observed. a.u., arbitrary units. **c**, Schematic illustration of the (interleaved) randomized benchmarking sequence applied to Q1.  $C$  corresponds to a single Clifford gate, with  $m$  being the total number of applied random Clifford gates. **d**, Differential bias current as a function of  $m$  for the randomized benchmarking sequence on Q1. The extracted control fidelity is  $F_c = (99.3 \pm 0.05)\%$ . **e**, Gate fidelities for the  $\pi$  and  $\pi/2$  gates. Error bars correspond to  $1\sigma$ . **f**, Spin coherence and life times for Q1 and Q2. Error bars correspond to  $1\sigma$ .

which can be mitigated to a large extent by using nuclear notch filtering<sup>29</sup>, leading to  $T_2 = 800$   $\mu\text{s}$ . This source of dephasing can be avoided altogether by using group IV materials with nuclear spin-free isotopes<sup>30</sup>. This has led to  $T_2 = 28$  ms for electrons in isotopically purified silicon<sup>13</sup>, and isotopic purification may also increase the quantum coherence in germanium. Furthermore, we observe spin lifetimes of  $T_{1,Q1} = 9$   $\mu\text{s}$  and  $T_{1,Q2} = 3$   $\mu\text{s}$ . We have found that these lifetimes increase exponentially when lowering the tunnel coupling between each qubit and its respective reservoir (Extended Data Fig. 8), and relaxation times of  $T_1 > 100$   $\mu\text{s}$

have been reported for germanium nanowires<sup>19,20</sup>, both giving good prospects for increasing the relaxation time by closing the reservoir barrier during operation.

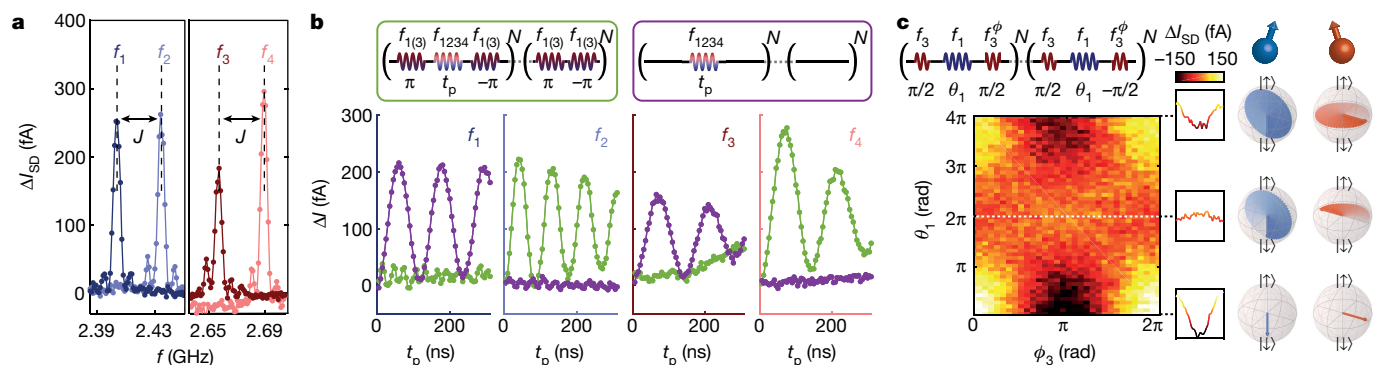
When the manipulation of both qubits is combined, the coupling of the two qubits (exchange interaction $J$ ) becomes apparent. As is illustrated in Fig. 3a, the resonance frequency of each of the qubits is shifted when the other qubit is prepared in its  $|\uparrow\rangle$  state. The strength of this interaction depends on the inter-dot tunnel coupling  $t_{12}$  as well as the detuning  $\varepsilon$  of the dot potentials. By changing the amplitude of voltage



**Fig. 3 | Tunable exchange coupling and operation at the charge symmetry point.** **a**, Illustration of the relevant energy levels in our hole double quantum dot with zero (green) and finite (black) exchange coupling  $J$  between the dots. Six energy levels are considered: the four different  $(1,1)$ -charge states as well as the  $(2,0)_s$  and  $(0,2)_s$  singlet charge states in which both holes occupy the same quantum dot. Four individual transitions can be driven, corresponding to the conditional rotations of the two-qubit system. The size of the exchange interaction is equal to  $J/h = f_2 - f_1 = f_4 - f_3$ . **b**, Measurement pulse cycles used to map out the exchange splitting of Q1 (top) and Q2 (bottom). As a result of the demodulation of the alternating cycles, transition  $f_{1,(3)}$  gives a negative signal and transition  $f_{2,(4)}$  results in a positive signal. **c**, **d**, EDSR spectra of Q1 (**c**) and Q2 (**d**) as a function of the detuning  $\varepsilon$ . The exchange splitting can be tuned to a minimum at  $\varepsilon = 0$  and increases closer to the  $(m, n) - (m+1, n-1)$  and  $(m, n) - (m-1, n+1)$  charge transitions. **e**, Exchange interaction as a function of  $\varepsilon$  as

extracted from **c**, **d**. Fitting the exchange coupling yields an interdot tunnel coupling  $t_{12}/h = 1.8$  GHz and charging energy  $U = 1.46$  meV. **f**, The interdot tunnel coupling can also be controlled by gate BC. Changing the potential on this gate, while keeping  $\varepsilon = 0$ , allows good control over the exchange interaction between the two qubits. **g**, Coherence time  $T_2^*$  of both qubits as a function of detuning voltage  $V_\varepsilon$ . When the slope of the resonance line is equal to zero, the qubit is expected to be, to first order, insensitive to charge noise. Solid lines indicate fits of the data to  $(a \frac{\partial f_{\text{res}}}{\partial V_\varepsilon} + T_0)^{-1}$ , with  $\frac{\partial f_{\text{res}}}{\partial V_\varepsilon}$  the numerical derivative of the resonance line frequency as a function of detuning,  $T_0$  the residual decoherence and  $a$  a scaling factor. It can be observed that  $T_2^*$  is indeed longest when the slope of the resonance line is closest to zero. Error bars correspond to  $1\sigma$ . **h**, Resonance frequency of transition  $f_1$  and  $f_3$  as a function of detuning voltage.





**Fig. 4 | Fast two-qubit logic with germanium qubits.** **a**, EDSR spectra of both qubits. Resonance peaks can be observed, corresponding to the four individual transitions indicated in Fig. 3a. The peaks are power-broadened, and the linewidth is thus determined by the Rabi frequency. **b**, Controlled qubit rotations can naturally be performed by selectively driving each of the four transitions. A CX gate is achieved at  $t_{\text{CX}} = t_{\pi}$  on  $f_1$  ( $f_3$ ). A small off-resonant

driving effect can be observed, which we mitigate by tuning  $t_{\text{CX}} = t_{\pi, \text{resonant}} = t_{4\pi, \text{off-resonant}}$ . **c**, Colour plot of  $\Delta I_{\text{SD}}$  as a function of Q1 CX-pulse length,  $\theta_1$ , and the phase of the second  $\pi/2$ -rotation on Q2,  $\phi_3$ . Owing to the  $Z(\theta_1/2)$  rotation on the control qubit, a  $\pi$  phase shift can be observed on Q2 for a conditional  $2\pi$  rotation on Q1 ( $f_1$ ).

pulse to point M (dotted line in Fig. 1d), we can map  $J$  as a function of the detuning  $\varepsilon$ . This is shown in Fig. 3c, d, where the subtraction of two pulse sequences in the measurement (see Fig. 3b) results in a positive signal for the unprepared qubit resonances and a negative signal for the prepared states (see Extended Data Fig. 2). As shown in Fig. 3e, the exchange coupling that is reflected in the frequency difference between the initial and prepared resonance positions, is very well described by a simple model<sup>2,31</sup> using  $J = 4Ut_{12}^2/[U^2 - (\alpha\varepsilon - U_0)^2]$ . Here,  $U$  is the charging energy of the quantum dots,  $\alpha = 0.23$  is the lever arm of P1 and P2, and the interdot tunnel coupling is  $t_{12}/h = 1.8$  GHz. In addition, the strength of  $t_{12}$  can be tuned by using the central barrier BC (Fig. 3f). Here, we use a virtual gate voltage<sup>3</sup>  $V_{\text{t12}}$ , where  $V_{\text{BC}}$  is set while compensating its influence on the dot potentials by appropriate corrections to  $V_{\text{P1}}$  and  $V_{\text{P2}}$ . As a result of this full control over the coupling, we are able to operate the qubits at a mostly charge-insensitive point of symmetric detuning<sup>4,5</sup>, where the qubit resonance frequencies are the least susceptible to changes in the electric field, while choosing an exchange coupling strength large enough for rapid two-qubit operations. The advantage of this reduced sensitivity to detuning noise is demonstrated in Fig. 3g, where the dephasing time  $T_2^*$  of both qubits is measured as a function of  $\varepsilon$ . Here,  $T_2^*$  strongly increases where the slope of  $f_{1(3)}$  with respect to the detuning for Q1 (Q2) is minimal, with the longest average phase coherence reached in the flat region at  $V_{\varepsilon} \approx 6$  mV.

The direct control over the tunnel coupling enables us to tune the exchange interaction to a sizable strength of  $J/h = 39$  MHz at the symmetry point, as demonstrated in Fig. 4a. We exploit this to obtain fast selective driving and operate in an exchange always-on regime<sup>8,32</sup>. Full control is obtained by applying microwave pulses at the four resonant frequencies, while further gate pulses controlling  $J$  are not needed. A pulse at a single resonance frequency will result in a conditional rotation of the target qubit, as we show in Fig. 4b. A CX-operation can be achieved by setting  $t_{\text{CX}} = t_{\pi, \text{resonant}} = t_{4\pi, \text{off-resonant}}$ . A fast CX-operation is thus achieved within  $t_{\text{CX}, \text{Q1}} = 55$  ns and  $t_{\text{CX}, \text{Q2}} = 75$  ns, with Q1 and Q2 as the target qubits respectively.

As a result of the pulsing, we observe a minor shift in the resonance frequency of both qubits, observed before in Si/SiGe quantum dots<sup>33</sup>. We compensate the temporary change in resonance frequency by applying phase corrections to all following pulses (see Extended Data Fig. 9). In Fig. 4c, we show the effect of a controlled rotation on the

control qubit with applied phase corrections. We observe a larger signal amplitude on Q1 after 0 and  $4\pi$  rotations on Q2 as compared with a  $2\pi$  rotation on Q2. This  $4\pi$  periodicity is in agreement with fermionic statistics and suggests an echoing pulse correcting residual environmental coupling. The full  $\pi$  phase shift on Q2 for a conditional  $2\pi$  rotation on Q1, as a result of the  $\theta_1/2$  phase that is accumulated by the control qubit, demonstrates the application of a coherent CX gate.

The demonstration of a universal gate set with all-electrical control and without the need of any microscopic structures offers good prospects to scale up spin qubits using holes in strained germanium. The hole states do not suffer from nearby valley states, and the quantum dots are contacted by superconductors<sup>22</sup> that may be shaped into microwave resonators for spin-photon coupling. This provides opportunities for a platform that can combine semiconducting, superconducting and topological systems for hybrid technology with fast and coherent control over individual hole spins. Moreover, the demonstrated quantum coherence and level of control make planar germanium a natural candidate to engineer artificial Hamiltonians for quantum simulation, going beyond classically tractable experiments.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-019-1919-3>.

- Ladd, T. D. et al. Quantum computers. *Nature* **464**, 45–53 (2010).
- Loss, D. & DiVincenzo, D. P. Quantum computation with quantum dots. *Phys. Rev. A* **57**, 120–126 (1998).
- Hensgens, T. et al. Quantum simulation of a Fermi-Hubbard model using a semiconductor quantum dot array. *Nature* **548**, 70–73 (2017).
- Reed, M. et al. Reduced sensitivity to charge noise in semiconductor spin qubits via symmetric operation. *Phys. Rev. Lett.* **116**, 110402 (2016).
- Martins, F. et al. Noise suppression using symmetric exchange gates in spin qubits. *Phys. Rev. Lett.* **116**, 116801 (2016).
- Veldhorst, M. et al. A two-qubit logic gate in silicon. *Nature* **526**, 410–414 (2015).
- Yoneda, J. et al. A quantum-dot spin qubit with coherence limited by charge noise and fidelity higher than 99.9%. *Nat. Nanotechnol.* **13**, 102–106 (2017).
- Huang, W. et al. Fidelity benchmarks for two-qubit gates in silicon. *Nature* **569**, 532–536 (2019).
- Zajac, D. M. et al. Resonantly driven CNOT gate for electron spins. *Science* **359**, 439–442 (2018).
- Koppens, F. H. L. et al. Driven coherent oscillations of a single electron spin in a quantum dot. *Nature* **442**, 766–771 (2006).

11. Petta, J. R. *et al.* Coherent manipulation of coupled electron spins in semiconductor quantum dots. *Science* **309**, 2180–2184 (2005).
12. Foletti, S., Bluhm, H., Mahalu, D., Umansky, V. & Yacoby, A. Universal quantum control of two-electron spin quantum bits using dynamic nuclear polarization. *Nat. Phys.* **5**, 903–908 (2009).
13. Veldhorst, M. *et al.* An addressable quantum dot qubit with fault-tolerant control-fidelity. *Nat. Nanotechnol.* **9**, 981–985 (2014).
14. Bulaev, D. V. & Loss, D. Spin relaxation and decoherence of holes in quantum dots. *Phys. Rev. Lett.* **95**, 076805 (2005).
15. Bulaev, D. V. & Loss, D. Electric dipole spin resonance for heavy holes in quantum dots. *Phys. Rev. Lett.* **98**, 097202 (2007).
16. Maurand, R. *et al.* A CMOS silicon spin qubit. *Nat. Commun.* **7**, 13575 (2016).
17. Watzinger, H. *et al.* A germanium hole spin qubit. *Nat. Commun.* **9**, 3902 (2018).
18. Liles, S. D. *et al.* Spin and orbital structure of the first six holes in a silicon metal-oxide-semiconductor quantum dot. *Nat. Commun.* **9**, 3255 (2018).
19. Hu, Y., Kuemmeth, F., Lieber, C. M. & Marcus, C. M. Hole spin relaxation in Ge–Si core-shell nanowire qubits. *Nat. Nanotechnol.* **7**, 47–50 (2012).
20. Vukušić, L. *et al.* Single-shot readout of hole spins in Ge. *Nano Lett.* **18**, 7141–7145 (2018).
21. Dobbie, A. *et al.* Ultra-high hole mobility exceeding one million in a strained germanium quantum well. *Appl. Phys. Lett.* **101**, 172108 (2012).
22. Hendrickx, N. W. *et al.* Gate-controlled quantum dots and superconductivity in planar germanium. *Nat. Commun.* **9**, 2835 (2018).
23. Sammak, A. *et al.* Shallow and undoped germanium quantum wells: a playground for spin and hybrid quantum technology. *Adv. Funct. Mater.* **29**, 1807613 (2019).
24. Lodari, M. *et al.* Light effective hole mass in undoped Ge/SiGe quantum wells. *Phys. Rev. B* **100**, 041304 (2019).
25. Nenashev, A. V., Dvurechenskii, A. V. & Zinovieva, A. F. Wave functions and g factor of holes in Ge/Si quantum dots. *Phys. Rev. B* **67**, 205301 (2003).
26. Maier, F., Kloeffel, C. & Loss, D. Tunable g factor and phonon-mediated hole spin relaxation in Ge/Si nanowire quantum dots. *Phys. Rev. B* **87**, 161305 (2013).
27. Knill, E. *et al.* Randomized benchmarking of quantum gates. *Phys. Rev. A* **77**, 012307 (2008).
28. Hutin, L. *et al.* in *2018 48th European Solid-State Device Research Conference (ESSDERC)*, 12–17 (2018).
29. Malinowski, F. K. *et al.* Notch filtering the nuclear environment of a spin qubit. *Nat. Nanotechnol.* **12**, 16–20 (2017).
30. Itoh, K. M. & Watanabe, H. Isotope engineering of silicon and diamond for quantum computing and sensing applications. *MRS Commun.* **4**, 143–157 (2014).
31. Russ, M. *et al.* High-fidelity quantum gates in Si/SiGe double quantum dots. *Phys. Rev. B* **97**, 085421 (2018).
32. Vandersypen, L. M. K. & Chuang, I. L. NMR techniques for quantum control and computation. *Rev. Mod. Phys.* **76**, 1037–1069 (2005).
33. Takeda, K. *et al.* Optimized electrical control of a Si/SiGe spin qubit in the presence of an induced frequency shift. *npj Quantum Inf.* **4**, 54 (2018).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2020

### Fabrication process

Our Ge/SiGe heterostructures are grown on a 100-mm n-type Si(001) substrate, using an Epsilon 2000 (ASMI) RP-CVD reactor, as described in ref.<sup>23</sup>. The device's Ohmic contacts and the electrostatic gates are defined by electron beam lithography, electron beam evaporation and lift-off of Al and Ti/Pd. Ohmic contacts consist of a 20-nm-thick Al layer, followed by a 17-nm-thick Al<sub>2</sub>O<sub>3</sub> gate dielectric grown by atomic layer deposition at 300 °C. Next, the first layer of Ti/Pd (40 nm) gates is deposited, followed by 17 nm of Al<sub>2</sub>O<sub>3</sub> and the second layer of overlapping Ti/Pd (40 nm) gates. Finally, vias contacting the lower gate layer are etched through the top Al<sub>2</sub>O<sub>3</sub> layer, followed by the deposition of 1-μm-thick Al<sub>99</sub>Si<sub>1</sub> bond pads to protect the device during bonding.

### Experimental set-up

All measurements are performed in a Bluefors dry dilution refrigerator with a base temperature of  $T_{\text{bath}} \approx 10$  mK. Constant d.c. voltages are applied with battery-powered voltage sources, and the voltages on gates P1 and P2 are combined with an a.c. voltage by a bias-tee with a cut-off frequency of 3 Hz. The a.c. voltage for gate P1 is generated by an arbitrary waveform generator (AWG) Tektronix AWG5014C, combined with a microwave signal generated by a Keysight PSG8267D vector source. The a.c. voltage for gate P2 is solely the waveform generated by the AWG. EDSR pulses are generated by the PSG8267D using the internal IQ-mixer, driven by two output channels of the AWG. Both qubits can be addressed by setting the vector source to an intermediate frequency of typically  $f_{\text{PSG}} = 2.56$  GHz, and IQ-mixing this with a (co)sine wave generated on channels 3 and 4 of the AWG. Because the on/off ratio of the IQ-modulation of our vector source is only 40 dB and small residual output power may lead to added infidelity, we use digital pulse modulation in series with the IQ modulation. The pulse modulation is driven by the AWG and is turned on 15 ns before the first pulse and turned off 7 ns after the last pulse in the sequence, resulting in a total suppression of 120 dB when the source is off.

We typically apply a source-drain bias voltage of  $V_{\text{SD}} = 0.3$  mV and measure the current through the device using an in-house-built transimpedance amplifier, after which the signal is low-pass filtered at 10 kHz and measured using an Stanford Research SR830 lock-in amplifier, as described in Methods section 'Sequence details' below.

### Virtual gates

To allow independent control over the tunnel coupling and the charge occupation of the double dot system, we make use of virtual gates<sup>3</sup>. When changing the different barrier gate voltages, linear corrections are applied to the device's plunger gates to correct for the cross-capacitance between the different gates. These coefficients are obtained from the relative slopes of the charge-addition lines with respect to the different device gates and normalized to the respective plunger gate coefficient. We write

$$\begin{pmatrix} \text{VP1} \\ \text{VP2} \end{pmatrix} = \begin{pmatrix} \alpha_{\text{P1,P1}} & \alpha_{\text{P2,P1}} & \alpha_{\text{BC,P1}} & \alpha_{\text{BR1,P1}} & \alpha_{\text{BR2,P1}} \\ \alpha_{\text{P1,P2}} & \alpha_{\text{P2,P2}} & \alpha_{\text{BC,P2}} & \alpha_{\text{BR1,P2}} & \alpha_{\text{BR2,P2}} \end{pmatrix} \delta \begin{pmatrix} \text{P1} \\ \text{P2} \\ \text{BC} \\ \text{BR1} \\ \text{BR2} \end{pmatrix}$$

with VP1 and VP2 the virtual plunger gates, and P1, P2, BC, BR1 and BR2 the different physical device gates as indicated in Fig. 1a. The virtual gate matrix describes the different couplings and is given by

$$\begin{pmatrix} \alpha_{\text{P1,P1}} & \alpha_{\text{P2,P1}} & \alpha_{\text{BC,P1}} & \alpha_{\text{BR1,P1}} & \alpha_{\text{BR2,P1}} \\ \alpha_{\text{P1,P2}} & \alpha_{\text{P2,P2}} & \alpha_{\text{BC,P2}} & \alpha_{\text{BR1,P2}} & \alpha_{\text{BR2,P2}} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0.8 & 0.35 & 0 \\ 0 & 1 & 0.8 & 0 & 0.4 \end{pmatrix}$$

We do not correct for the crosstalk between the two plunger gates, such that  $\alpha_{\text{P2,P1}} = \alpha_{\text{P1,P2}} = 0$ . The crosstalk between the quantum dot and the reservoir barrier of the other dot is negligible because of their physical separation. Furthermore, it can be observed that the coupling

of the centre barrier to both dots is approximately twice as strong as the reservoir barriers as a direct effect of its increased size.

### Sequence details

To improve the quality of the transport measurements, we establish a lock-in measurement scheme in which the measurement of interest is alternated with a reference measurement to account for slow variations in the transport current through the device, as well as temperature-dependent drifts in our transimpedance amplifier, as is illustrated in Extended Data Fig. 1. The measurement cycle, consisting of the readout as well as the manipulation phase, typically has a length of  $\tau_{\text{MC}} \approx 1$  μs. With the AWG, we generate a waveform that repeats the measurement cycle  $N$  times, followed by  $N$  repetitions of a similar reference measurement, with  $N$  chosen such that these cycles alternate at a lock-in frequency of  $f_{\text{lock-in}} = 89.75$  Hz. The measured transport current is then demodulated by a lock-in amplifier, using a reference signal generated by the AWG. As a result, the lock-in output signal will be directly related to the difference in transport current between the measurement and the reference cycle. During the readout, no differential current is observed when the qubits are in their  $|\downarrow\downarrow\rangle$  ground state, whereas a signal of typically  $\Delta I_{\text{SD}} \approx 0.3$  pA is measured for all other spin configurations and a total cycle length of  $t_{\text{cycle}} = 900$  ns. This is in good agreement with a bias current  $\Delta I = 2e/t_{\text{cycle}} = 0.4$  pA, as expected for the random loading of a hole spin.

For a Rabi experiment, the measurement cycle contains a single microwave pulse of duration  $t_p$ , whereas the reference cycle has no pulses. In the case of a Ramsey experiment, both the measurement and reference cycle contain a  $\pi/2$  pulse, a wait  $\tau$  and a final  $\pi/2$  pulse, but in the reference cycle the final  $\pi/2$  pulse is phase-shifted by  $\phi = \pi$ . This will result in an opposite projection for the two measurements and thereby maximum differential signal. For the randomized benchmarking, a similar scheme is used (see Fig. 2a), where the recovery pulse in the measurement cycle is chosen to project to the spin-up state, while the recovery pulse in the reference cycle projects to the spin-down state, resulting in an exponential decay towards  $\Delta I_{\text{SD}} = 0$ . Each data point is averaged over approximately  $10^5$  repetitions of 1,500 randomly drawn gate sequences. Finally, for the exchange measurements, we alternate a measurement cycle where we apply a  $\pi$  and  $-\pi$  pulse to Q1 (Q2) before and after the probing pulse respectively, with a reference cycle where Q1 (Q2) is not pulsed. When the probing pulse is off-resonant with both resonance frequencies, the measurement cycle gives effectively no rotation of Q1 (Q2) and the reference cycle does not result in any rotation. As a result the demodulated signal will be zero. When the probing pulse frequency is on resonance with the unprepared resonance frequency  $f_3$  ( $f_1$ ), the measurement cycle will still be an effective zero rotation on Q1 (Q2) due to the selective driving of  $f_3$  ( $f_1$ ) and thus give no signal. The reference cycle will now result in a  $\pi$  rotation on Q2 (Q1) and will therefore give a signal, resulting in a negative demodulated signal. In the case where the probing pulse is resonant with the prepared resonance line  $f_4$  ( $f_2$ ), the measurement cycle will generate a signal whereas the reference cycle will give no signal, thus resulting in a positive demodulated signal. All different pulse cycle configurations and the respective qubit projections are illustrated in Extended Data Fig. 2b.

### Phase corrections for pulsing

We observe a shift of the resonance frequency of the qubits as a function of the microwave driving power. We attribute this to a rectification of the microwave signal, resulting in a d.c. voltage pulse which can modulate the resonance frequency through the SOC and exchange interaction. As a result of the shift during the pulsing, each qubit picks up a phase when it is idling, as well as an additional phase due to the pulses on the other qubit. We can calibrate these frequency shifts and correct all following pulses to counteract this phase shift.

To probe the effect of all possible pulses on all possible resonances, we perform an extended Ramsey experiment. We prepare a pulse sequence



consisting of two  $\pi/2$  pulses with a test gate (each of the four resonance lines, as well as idling) and  $\pi$  phase-shifted test gate in between, as indicated in Extended Data Fig. 9. For the experiment on  $f_2$  and  $f_4$ , we add an additional preparation and projection pulse at the start and end respectively, as indicated in grey in Extended Data Fig. 9. The back-and-forth rotation on the test gate cancels any driving effects, as well as the  $\theta/2$  phase picked up due to the conditional rotation, and leaves us with only the detuning phase. We now plot the transport current  $\Delta I_{SD}$  as a function of the phase  $\phi$  of the second  $\pi/2$ -pulse, as well as the length of the test gate. As a result of the frequency shift caused by the test gate, we observe a phase shift increasing linearly with the length of the test gate. We fit this phase shift for each gate, and we apply a correction to all following gates. Extended Data Fig. 9 shows the phase evolution for all test gates on all four resonance lines, both without corrections (Extended Data Fig. 9a), as well as with corrections applied (Extended Data Fig. 9b).

### Data availability

All data underlying this study are available from the 4TU ResearchData repository at <https://doi.org/10.4121/uuid:95bc1f2e-0218-4c55-8e5b-2b59e8fcc5e6>.

34. He, L., Bester, G. & Zunger, A. Electronic phase diagrams of carriers in self-assembled quantum dots: violation of Hund's rule and the Aufbau principle for holes. *Phys. Rev. Lett.* **95**, 246804 (2005).
35. Reuter, D. et al. Coulomb-interaction-induced incomplete shell filling in the hole system of InAs quantum dots. *Phys. Rev. Lett.* **94**, 026808 (2005).
36. Hensen, B. et al. A silicon quantum-dot-coupled nuclear spin qubit. *Nat. Nanotechnol.* Preprint at <http://arxiv.org/abs/1904.08260> (2019).
37. Crippa, A. et al. Electrical spin driving by  $g$ -matrix modulation in spin-orbit qubits. *Phys. Rev. Lett.* **120**, 137702 (2018).

**Acknowledgements** We thank L. M. K. Vandersypen, S. Dobrovitski and J. Helsen for valuable discussions. We acknowledge support through a FOM Projectruimte of the Foundation for Fundamental Research on Matter (FOM), associated with the Netherlands Organisation for Scientific Research (NWO).

**Author contributions** N.W.H. and D.P.F. performed the experiments. N.W.H. fabricated the device. A.S. and G.S. supplied the heterostructures. N.W.H., D.P.F. and M.V. wrote the manuscript with the input of all other authors. M.V. conceived and supervised the project.

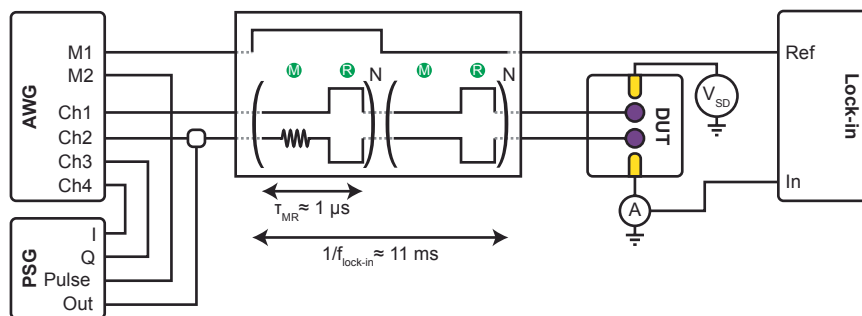
**Competing interests** The authors declare no competing interests.

### Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41586-019-1919-3>.

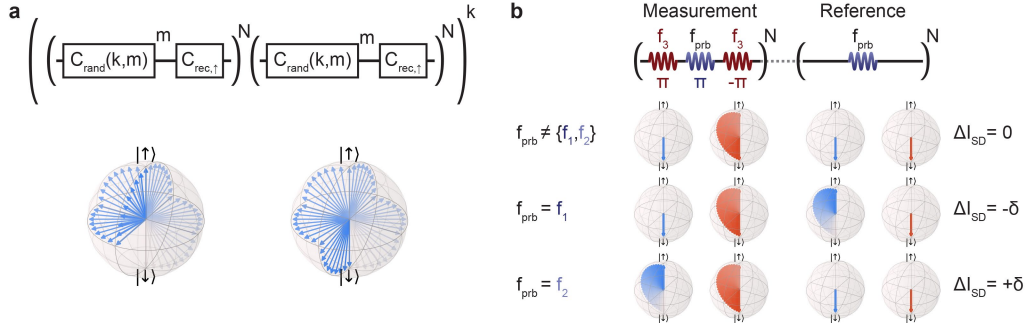
**Correspondence and requests for materials** should be addressed to M.V.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>.



**Extended Data Fig. 1 | Instrumentation set-up for the lock-in transport measurements.** Illustration of the set-up and relevant signals for the lock-in transport measurements. The AWG is used to generate alternating pulse cycles consisting of a repeated measurement and a repeated reference. The signal is

demodulated in a lock-in amplifier to give a direct measure of the difference between the two measurements and subtract slow variations in the transport signal.

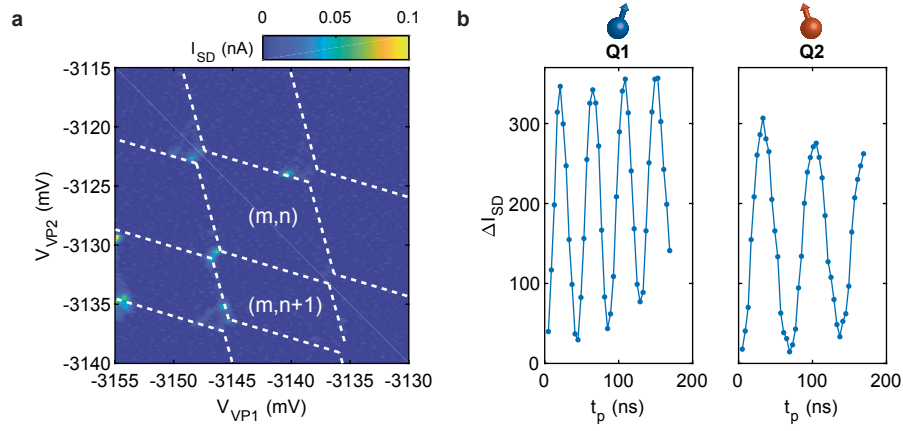


**Extended Data Fig. 2 | Pulse cycles used for the transport measurements.**

**a** Pulse cycles used for the randomized benchmarking experiments. The measurement pulse cycle consists of  $m$  gates randomly drawn from the Clifford group  $C_{\text{rand}}$  and a final Clifford gate projecting the qubit onto the spin-up state. The reference pulse cycle consists of the same  $m$  Clifford gates and a different final Clifford gate projecting the qubit onto the spin-down state. Each cycle is repeated  $N$  times, and a series of typically  $k = 50$  independent randomly drawn measurement and reference pulse cycles are alternated. These  $k = 50$  different draws are thus hardware-averaged on the lock-in amplifier, and the entire

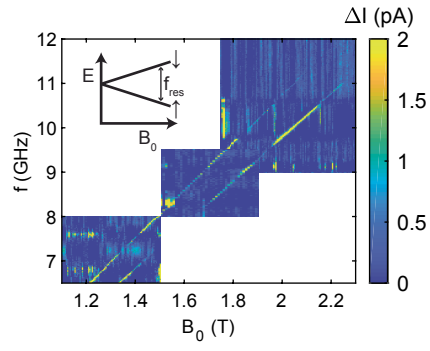
experiment is repeated and averaged 30 times, yielding a total approximate  $10^5$  repetitions of 1,500 different randomly drawn Clifford sequences of length  $m$ . An example of the qubit evolution for each pulse cycle is plotted on the Bloch sphere below. **b**, Pulse cycles used for the exchange mapping experiments. The measurement pulse cycle consists of a broad preparation and restoring pulse at frequency  $f_3$  ( $f_j$ ), around a probing pulse at frequency  $f_{\text{prb}}$ . The reference pulse cycle consists solely of the probing pulse at  $f_{\text{prb}}$ . The qubit evolutions for the different resonance conditions are plotted on the Bloch sphere and illustrate the different signals measured in Fig. 4c, d.



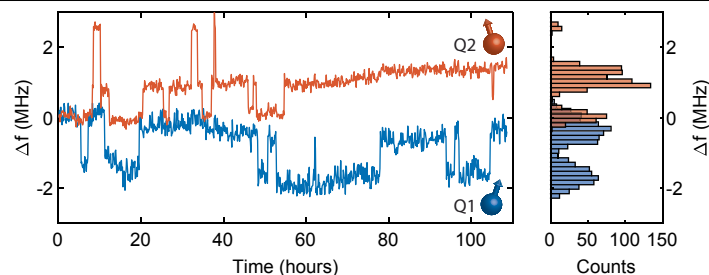


**Extended Data Fig. 3 | Demonstration of qubit operation at a second hole occupancy.** **a**, Charge stability diagram showing the  $(m, n)$  hole occupancy used during all experiments in the main text, as well as the  $(m, n+1)$  occupancy for which we observe PSB as well. For an unpolarized filling of the quantum dots, one expects an alternating suppression of the transport current due to PSB, as spin blockade occurs only when an orbital level is fully occupied. However, the spin-filling for holes is known to be highly polarized<sup>34,35</sup>, and

therefore PSB can occur in sequential quantum dot fillings. **b**, Coherent Rabi oscillations measured in the  $(m, n+1)$  occupancy. A slight linear offset is observed for Q1, which can be attributed to the microwave power. We note that, for the same microwave power, the Rabi frequency of Q2 in the  $(m, n+1)$  occupancy is increased substantially compared to the  $(m, n)$  filling. We attribute this to the hole being in a different orbital, where the effective SOC may be different.



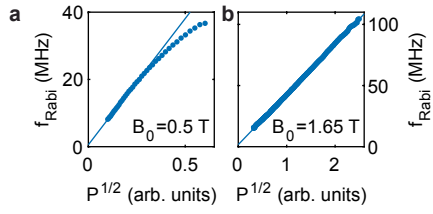
**Extended Data Fig. 4 | Qubit resonance frequencies as a function of magnetic field.** Colour plot indicating the transport current  $\Delta I$  through the double dot system, as a function of external magnetic field  $B_0$  and the frequency  $f$  of the applied microwave signal. We have numerically subtracted the mean of each row and column in each of the three individual colour plots, to account for the slow drifts in transport current, as well as the line resonances in our fridge cabling. The two bright lines indicate an increase in the transport current due to the microwave rotating either spin and thus lifting PSB.

**Extended Data Fig. 5 | Temporal dependence of the resonance frequency.**

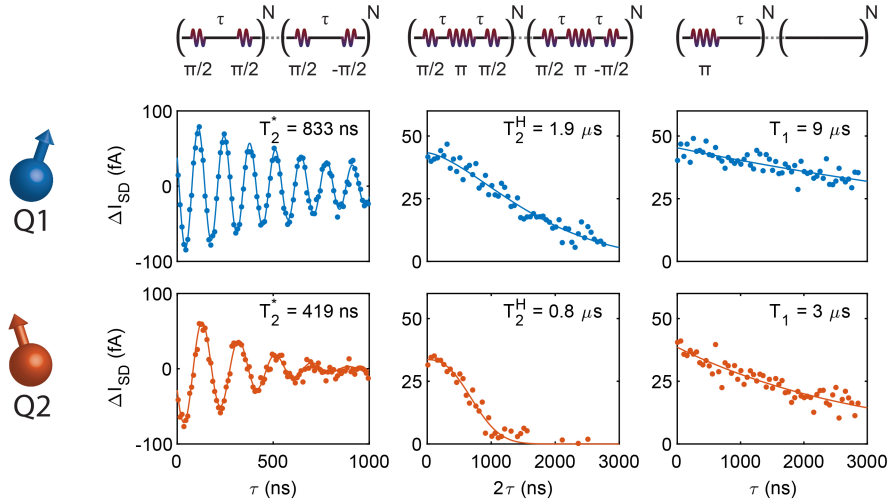
We track the resonance frequency of both Q1 and Q2 over the time of approximately 110 h. We observe that the qubit frequency remains remarkably stable over this period, but do observe discrete, uncorrelated steps in the resonance frequencies of both qubits. The resonance frequency of Q1 only shows steps of  $\Delta f \approx 2$  MHz between two distinct levels, whereas for Q2 we

observe steps of both  $\Delta f \approx 1$  MHz and  $\Delta f \approx 2$  MHz, between three different levels, as also becomes apparent from the histogram. The origin of these steps could be, for example, the slow loading and unloading of charge traps, which manipulates the qubit resonance frequency through the change in electric field, or hyperfine coupling to a nearby nuclear spin<sup>36</sup>.



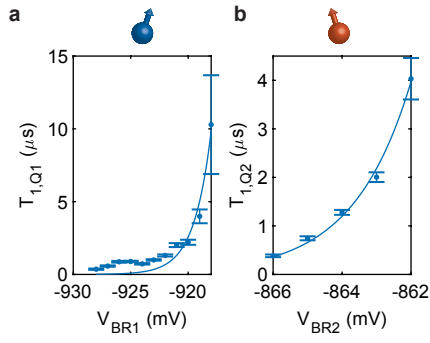


**Extended Data Fig. 6 | Magnetic field dependence of the driving speed of Q1.**  
**a, b,** Rabi frequency dependence on the applied microwave power  $P$  in arbitrary units, for  $B_0 = 0.5$  T (**a**) and  $B_0 = 1.65$  T (**b**). Multiple mechanisms can be at play for the EDSR driving of the spins<sup>37</sup> and these are typically all linearly dependent on  $B_0$ . As a result of this, considerably higher driving frequencies can be reached at higher magnetic fields. We note that the exact microwave power cannot be compared between the two measurements, owing to the strong frequency dependence of the attenuation of our fridge lines.



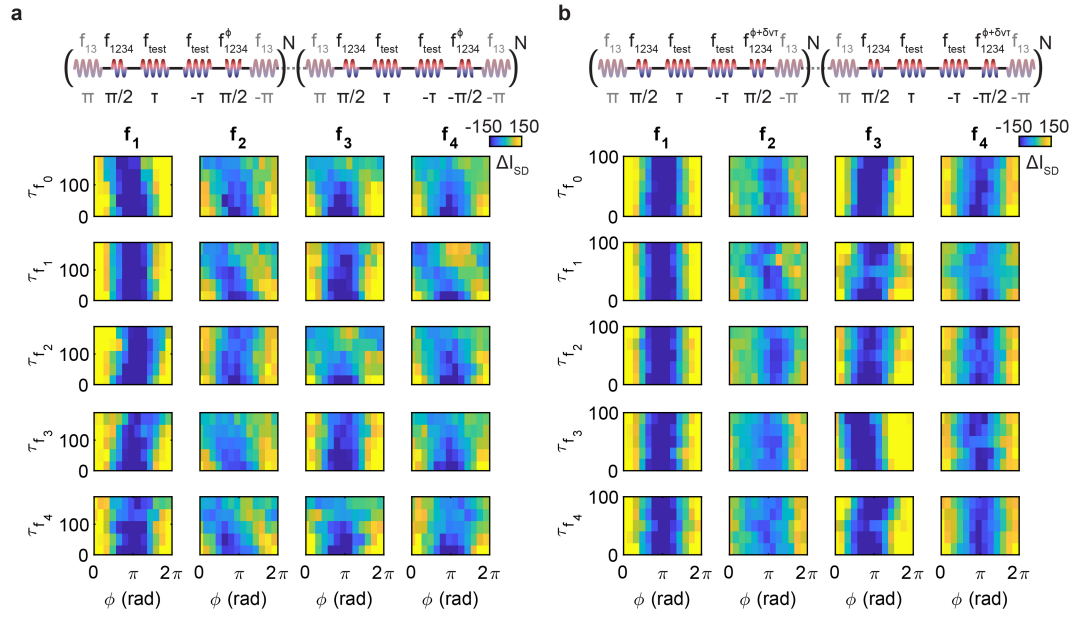
**Extended Data Fig. 7 | Relaxation, dephasing and coherence times.** We perform a Ramsey experiment, in which two  $\pi/2$  pulses are separated by time  $\tau$ , during which the qubit will evolve as a result of the implemented detuning. We fit the decay of the observed oscillations to  $\Delta I_{SD} = a \cos(2\pi\Delta f\tau + \phi) \exp[-(\tau/T_2^*)^{\alpha^*}]$ , with  $a$  a scaling factor,  $\Delta f$  the detuning and  $\phi$  a phase offset, and find a spin coherence time of  $T_{2,Q1}^* = 833$  ns and  $T_{2,Q2}^* = 419$  ns and decay coefficients of  $\alpha_{Q1}^* = 1.2 \pm 0.2$  and  $\alpha_{Q2}^* = 1.5 \pm 0.2$ , for Q1 and Q2, respectively. The spin coherence can be extended by performing a Hahn echoing sequence, consisting of  $\pi/2, \pi$

and  $\pi/2$  pulses separated by waiting times  $\tau$ . Fitting the observed decay as a function of the total waiting time  $2\tau$  to a power law  $\Delta I_{SD} = a \exp[-(2\tau/T_2^H)^{\alpha^H}]$ , we find extended coherence times of  $T_{2,Q1}^H = 1.9$   $\mu$ s and  $T_{2,Q2}^H = 0.8$   $\mu$ s and decay coefficients of  $\alpha_{Q1}^H = 1.5 \pm 0.1$  and  $\alpha_{Q2}^H = 2.5 \pm 0.3$ , for Q1 and Q2, respectively. Finally, we perform a measurement of the spin lifetime by applying a single  $\pi$  pulse, after which we wait for a time  $\tau$ . We fit the decay to  $\Delta I_{SD} = \exp[-(\tau/T_1)]$  and find lifetimes of  $T_{1,Q1} = 9$   $\mu$ s and  $T_{1,Q2} = 3$   $\mu$ s.



**Extended Data Fig. 8 | Relaxation time  $T_1$  as a function of gate voltage on the tunnel barriers between dot and reservoir. a, b,** The relaxation time  $T_1$  of the dots increases approximately exponentially as a function of the respective dot–reservoir gate voltage, for Q1 (a) as well as for Q2 (b). The relaxation time of Q1 increases exponentially from  $T_1 < 1 \mu\text{s}$  to  $T_1 > 10 \mu\text{s}$ , and a similar scaling is observed for Q2. For even smaller dot–reservoir couplings, the transport signal drops below our measurement limit, but switching to charge sensing could allow a further increase in  $T_1$ .





**Extended Data Fig. 9 | Phase corrections on the qubits. a**, Extended Ramsey experiment on each of the four resonance line, using five different test gates between the  $\pi/2$  pulses to observe the effect on the resonance frequency. A linear phase shift as a function of test gate pulse length  $\tau$  can be observed for

some lines, indicating a frequency shift during the pulsing. **b**, We compensate for this effect by performing a software update of  $\delta\phi = \delta\nu T$  to each additional pulse, with  $\delta\nu$  the frequency shift of the qubit as a result of the microwave signal.

# Conversion of non-van der Waals solids to 2D transition-metal chalcogenides

<https://doi.org/10.1038/s41586-019-1904-x>

Received: 2 March 2019

Accepted: 25 November 2019

Published online: 22 January 2020

Zhiguo Du<sup>1</sup>, Shubin Yang<sup>1\*</sup>, Songmei Li<sup>1</sup>, Jun Lou<sup>2</sup>, Shuqing Zhang<sup>3</sup>, Shuai Wang<sup>1</sup>, Bin Li<sup>1</sup>, Yongji Gong<sup>1</sup>, Li Song<sup>4</sup>, Xiaolong Zou<sup>3</sup> & Pulickel M. Ajayan<sup>2\*</sup>

Although two-dimensional (2D) atomic layers, such as transition-metal chalcogenides, have been widely synthesized using techniques such as exfoliation<sup>1–3</sup> and vapour-phase growth<sup>4,5</sup>, it is still challenging to obtain phase-controlled 2D structures<sup>6–8</sup>. Here we demonstrate an effective synthesis strategy via the progressive transformation of non-van der Waals (non-vdW) solids to 2D vdW transition-metal chalcogenide layers with identified 2H (trigonal prismatic)/1T (octahedral) phases. The transformation, achieved by exposing non-vdW solids to chalcogen vapours, can be controlled using the enthalpies and vapour pressures of the reaction products. Heteroatom-substituted (such as yttrium and phosphorus) transition-metal chalcogenides can also be synthesized in this way, thus enabling a generic synthesis approach to engineering phase-selected 2D transition-metal chalcogenide structures with good stability at high temperatures (up to 1,373 kelvin) and achieving high-throughput production of monolayers. We anticipate that these 2D transition-metal chalcogenides will have broad applications for electronics, catalysis and energy storage.

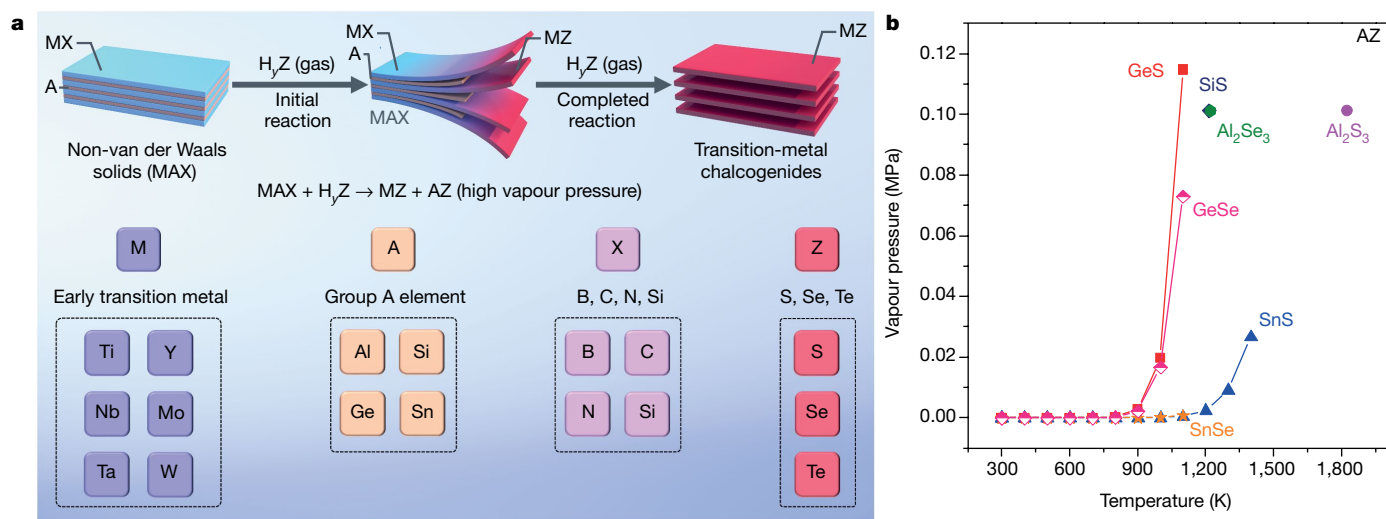
Two-dimensional (2D) atomic-layer crystals have demonstrated many unique physical and chemical properties as well as broad applications in electronics<sup>2</sup>, sensors<sup>5</sup>, catalysts<sup>9</sup> and batteries<sup>10,11</sup>. Generally, 2D structures such as graphene, boron nitride and transition-metal sulfides can be produced via a top-down approach, that is, by directly exfoliating the vdW counterparts through mechanical<sup>3</sup>, liquid-phase<sup>1</sup> and electrochemical procedures<sup>2</sup>. In this manner, various vdW materials—such as metal oxides<sup>12</sup>, hydroxides<sup>13</sup> and topological insulators<sup>14</sup>—can also be synthesized, enriching the 2D family of materials. In these 2D vdW nanocrystals, the elemental compositions, stoichiometric ratios and structural phases are usually inherited from their parent bulk counterparts, although 2D nanocrystals with phase-specific structures such as 1T and 2H phases are difficult to synthesize selectively<sup>6,7</sup>. Here we demonstrate an efficient topological conversion of non-vdW solids such as transition-metal carbides and nitrides under chalcogen vapours to 2D vdW transition-metal chalcogenide layers with identified 2H/1T phases, good stability at high temperatures (<1,373 K) and achieving high-throughput production of monolayers. We anticipate that the resultant transition-metal chalcogenide layers with favourable features would have broad applications for electronics, energy storage and conversions.

In the past decade, some unusual 2D nanocrystals have emerged from non-vdW solids such as haematite<sup>15</sup> or bulk layered transition-metal carbides and nitrides<sup>16</sup>, namely MAX phases, greatly increasing the number of 2D material compositions accessible. In particular, the non-vdW MAX phases—where M represents a transition-metal element, A usually represents an element from groups 13–16 of the periodic table and X is carbon or nitrogen—have predominantly mixed covalent or ionic

M–X bonds and metallic M–A bonds<sup>17,18</sup>. Because the M–A bonds are more chemically active than the M–X bonds, A species in MAX phases can be extracted using highly reactive solvents (hydrogen fluoride and strong bases)<sup>16,19</sup>, allowing few-layer-thick 2D transition-metal carbides, carbonitrides and nitrides—called MXenes—to be created. These 2D nanocrystals are usually terminated with defects and surface terminations of –OH, –O, –F or –Cl<sup>20–22</sup>. Owing to the very close atomic packing and strong chemical bonds in non-vdW solids, it remains a challenge to convert them to 2D nanocrystals with abundant exposed surfaces and identified phases.

Here we demonstrate an efficient strategy that enables us to convert a family of non-vdW bulk solids such as MAX phases to 2D transition-metal chalcogenides with well-defined phases. As depicted in Fig. 1a, under chalcogen-containing vapours (H<sub>2</sub>Z, where Z represents sulfur, selenium or tellurium and y is 0 or 2) at high temperatures, non-vdW MAX phases and transition-metal borides, silicides and carbides (Supplementary Fig. 1) have high activities. In particular, the active M–A bonds in MAX phases react easily with chalcogen-containing gases, resulting in products of AZ and MZ compositions. Such reactions must produce an AZ intermediate product at high vapour pressure, which would allow rapid evaporation rates, thus boosting the continuous reaction into the bulk of the reactant material. Thermodynamically, if the reaction temperature were high enough, all the post-transition-metal A (Si, Al, Sn, Ge) species in MAX phases could be transformed to metal chalcogenide gases (Supplementary Figs. 2–5), which facilitates the conversion of MAX phases to 2D nanostructures. As an example, based on temperature–vapour pressure relationships (Fig. 1b and Supplementary Fig. 6)<sup>23,24</sup>, germanium chalcogenides (GeS, GeSe)

<sup>1</sup>Key Laboratory of Aerospace Advanced Materials and Performance of Ministry of Education, School of Materials Science and Engineering, Beihang University, Beijing, China. <sup>2</sup>Department of Materials Science and NanoEngineering, Rice University, Houston, TX, USA. <sup>3</sup>Shenzhen Geim Graphene Center and Low-Dimensional Materials and Devices Laboratory (LDMD), Tsinghua-Berkeley Shenzhen Institute (TBSI), Tsinghua University, Shenzhen, China. <sup>4</sup>National Synchrotron Radiation Laboratory, CAS Center for Excellence in Nanoscience, University of Science and Technology of China, Hefei, China. \*e-mail: yangshubin@buaa.edu.cn; ajayan@rice.edu



**Fig. 1 | Schematic illustration of the conversion of non-vdW solids to 2D vdW transition-metal chalcogenides.** **a**, Non-vdW solids such as MAX phases are progressively transformed to 2D transition-metal chalcogenides via a topological conversion reaction ( $\text{MAX} + \text{H}_2\text{Z} (\text{gas}) \rightarrow \text{MZ} + \text{AZ}$ ), in which M

represents an early-transition-metal element, A is an element from groups 13–16, X is C, N, B or Si, and Z refers to S, Se and Te, associated with volatile AZ products. **b**, Temperature–vapour pressure relationships for various AZ substances.

have higher vapour pressures than other chalcogenide materials (SiS,  $\text{Al}_2\text{S}_3$ , SnS) at 1,073 K; from this we expect Ge-containing MAX phases to be easily converted to 2D transition-metal chalcogenides with vdW layers (Fig. 2c and Supplementary Fig. 7). According to the Clausius–Clapeyron equation<sup>25</sup> (see equation (1) in Methods), upon increasing the temperature further to more than 1,100 K, other Si-, Sn- and Al-containing MAX phases should also serve as precursors for the generation of 2D transition-metal chalcogenides (Supplementary Fig. 8), owing to the increased vapour pressures of the products at higher reaction temperatures. Using this principle, we have synthesized 13 transition-metal chalcogenides (Supplementary Tables 1 and 2), including 7 binary chalcogenides (based on the Ti-, Nb-, Mo- and Ta-containing MAX phases and MXenes), 5 heteroatom-doped chalcogenides with selected 2H phase or 1T phase and one composite (based on quaternary MAX phases). This demonstrates that our synthetic protocol is versatile, enabling the efficient conversion of a large number of non-vdW bulk solids to 2D transition-metal chalcogenides. Notably, although the resulting 2D structures are derived from bulk MAX phases, their compositions and stoichiometric ratios are very different from the parent compositions and are also different from other products commonly derived from MAX phases, such as MXenes<sup>16,26</sup>.

## 2D transition-metal chalcogenides (2H/1T)

As a proof of concept, we produced 2D transition-metal dichalcogenide (TMD)- $\text{MoS}_2$  nanocrystals via engineering MAX- $\text{Mo}_2\text{GeC}$  under hydrogen disulfide gas at 1,073 K (Fig. 2a; see Methods). X-ray diffraction patterns reveal the disappearance of  $\text{Mo}_2\text{GeC}$  peaks in the product (Fig. 2b). Instead, a series of diffraction peaks at  $14.1^\circ$ ,  $32.7^\circ$ ,  $39.5^\circ$  and  $58.3^\circ$  are well indexed to the (002), (100), (103) and (110) facets of hexagonal  $\text{MoS}_2$  (according to Joint Committee on Powder Diffraction Standards (JCPDS) Card No. 37-1492), demonstrating the complete conversion and removal of Ge-layers from MAX- $\text{Mo}_2\text{GeC}$  during our synthetic process. The resulting product exhibits uniform structure with largely extended spacing in the whole scanning electron microscope image (Fig. 2c and Supplementary Fig. 9), similar to those reported for expanded graphite and MXenes<sup>16</sup>. Transmission electron microscopy (TEM) (Fig. 2d and Supplementary Fig. 10) and high-resolution TEM (Fig. 2e) confirm clearly the highly exfoliated nanocrystals with a uniform interplanar

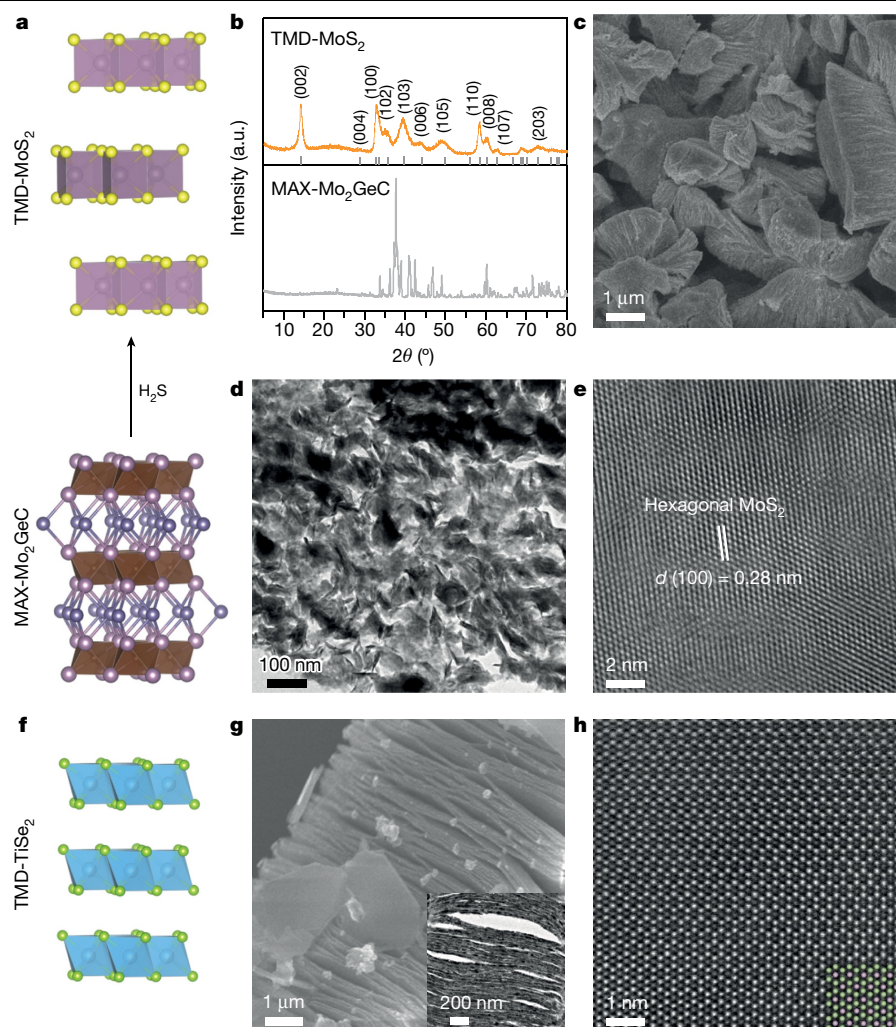
spacing of 0.28 nm, in good agreement with the spacing between the (100) facets of 2H  $\text{MoS}_2$  (ref. <sup>2</sup>).

To identify the interior structure of the resultant  $\text{MoS}_2$ , we conducted an ultrathin sectioning experiment. Many nanosheets with thicknesses from 0.4 nm to 4 nm are visible (Supplementary Fig. 11), indicating the co-existence of monolayers, bilayers and few layers in the sample. To inhibit the restacking of already-expanded 2D  $\text{MoS}_2$  and improve the fraction of monolayers, the converted samples were rapidly transferred to a low-temperature zone during our synthetic process. Thus, the fraction of monolayer  $\text{MoS}_2$  can be improved to 31% from 9% based on our standard conversion (Supplementary Figs. 12–16). Remarkably, when we directly converted thin non-vdW solid MXene- $\text{Mo}_2\text{CT}_x$  in  $\text{H}_2\text{S}$  gas, the fraction of monolayer  $\text{MoS}_2$  was up to about 91% (Supplementary Fig. 17). Raman spectra of the accordion-like  $\text{MoS}_2$  (Supplementary Figs. 18 and 19) show two typical peaks at  $379\text{ cm}^{-1}$  and  $405\text{ cm}^{-1}$ , corresponding to the in-plane  $\text{E}_{2g}^1$  and out-of-plane  $\text{A}_{1g}$  vibrational modes of 2H  $\text{MoS}_2$  (Supplementary Fig. 67)<sup>2</sup>, respectively. On the basis of thermodynamic considerations, by enhancing the reaction temperatures to more than 1,100 K, accordion-like  $\text{TiSe}_2$  could also be derived from MAX- $\text{Ti}_3\text{SiC}_2$  by substituting  $\text{H}_2\text{S}$  with selenium vapour (Fig. 2f–h, Supplementary Figs. 20–24), owing to the high vapour pressure of the SiSe product at such high temperatures (Supplementary Fig. 4). An atomic-resolution scanning transmission electron microscopy (STEM) image of  $\text{TiSe}_2$  (Fig. 2h) reveals the 1T superlattice with metal sites located at the centres of octahedral units. Such transformations suggest that our synthetic protocol can be generalized to convert non-vdW solids to vdW 2D nanocrystals with identified 2H/1T phases and high-throughput production of monolayers (Supplementary Figs. 25–36).

## 2D heteroatom-doped chalcogenides (2H/1T)

More than 70 ternary MAX phases<sup>27</sup> and some new quaternary MAX phases<sup>28</sup> such as  $(\text{W}_{2/3}\text{Y}_{1/3})_2\text{AlC}$ <sup>29</sup> and  $(\text{Ti}_{1/2}\text{Nb}_{1/2})_2\text{AlC}$ <sup>30</sup> have been explored, suggesting that it may be feasible to produce a series of transition-metal chalcogenides with multi-compositions via our topological conversion approach. One possibility is to produce accordion-like Y-doped  $\text{WS}_2$  with the 2H phase (Fig. 3a, b and Supplementary Figs. 37–39) based on a  $(\text{W}_{2/3}\text{Y}_{1/3})_2\text{AlC}$  precursor. After conversion with fast quenching, highly expanded accordion-like Y-doped  $\text{WS}_2$  can be obtained, in which the fraction of monolayers is up to 27% (Supplementary Figs. 40–45).





**Fig. 2 | Structural characterization of 2D transition-metal chalcogenides derived from MAX phases.** **a**, Schematic illustration of the conversion of MAX-Mo<sub>2</sub>GeC to accordion-like MoS<sub>2</sub> under H<sub>2</sub>S gas at 1,073 K. **b**, X-ray diffraction patterns of MAX-Mo<sub>2</sub>GeC and accordion-like MoS<sub>2</sub>. **c–e**, Scanning electron microscope (c), sectional TEM (d) and high-resolution TEM (e) images of

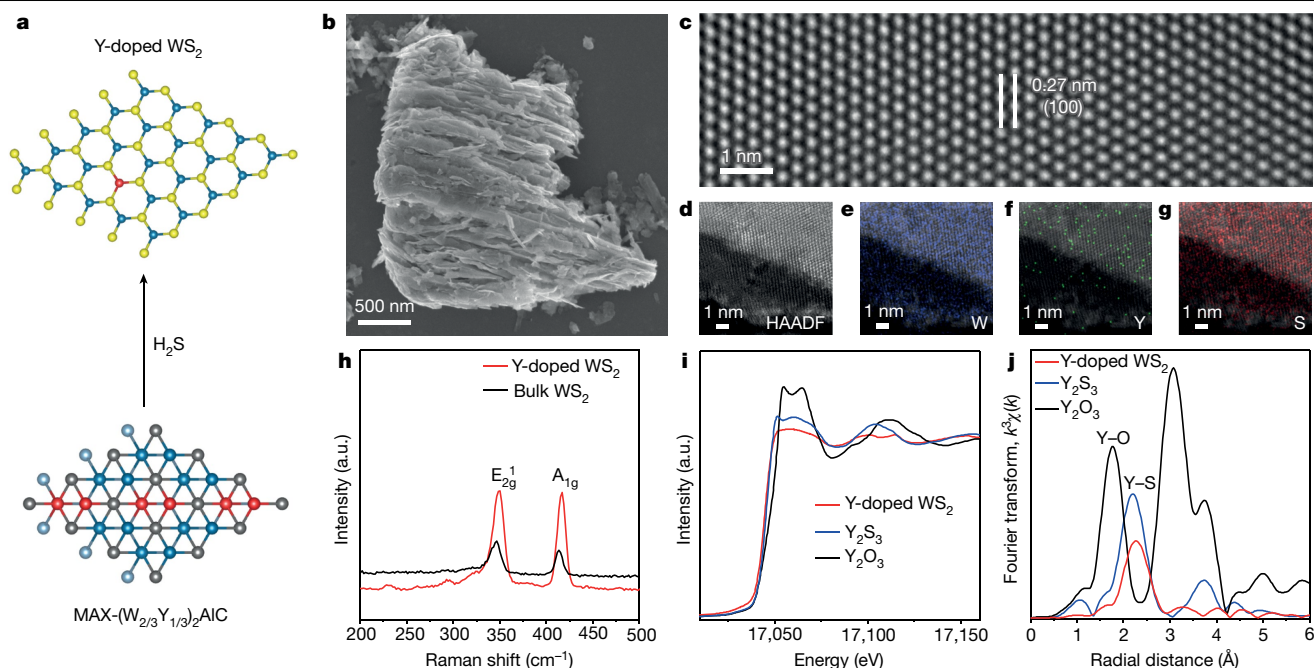
accordion-like MoS<sub>2</sub>. **f**, Crystalline structure of 1T TiSe<sub>2</sub>. **g**, Scanning electron microscope image and sectional TEM image (inset) of accordion-like TiSe<sub>2</sub>, indicating its highly expanded structure. **h**, Atomic-resolution STEM image of TiSe<sub>2</sub> layers and the corresponding atomic configuration (inset) of the 1T phase. Purple and green balls represent Ti and Se atoms, respectively.

High-resolution elemental mapping images (Fig. 3f) reveal that Y species are mainly monoatomically dispersed into the layers, associated with a few Y clusters, possibly accounting for the many Y–S bonds present (Supplementary Fig. 46). A typical 2H structure is visualized in an aberration-corrected TEM image (Fig. 3c), consistent with X-ray diffraction (Supplementary Fig. 38) and Raman analysis (Fig. 3h), indicating that the heteroatoms remain stable within the host 2D transition-metal chalcogenide structure. To further verify the Y–S bonds in Y-doped WS<sub>2</sub>, we conducted X-ray absorption near-edge fine structure (XANES) spectroscopy measurements. In the case of Y-doped WS<sub>2</sub>, the absorption edge in Y K-edge XANES spectra (Fig. 3i) is close to that of yttrium sulfide, suggesting that the Y is in the sulfide state. The Fourier transform spectra resulting from the analysis of Y-doped WS<sub>2</sub> by extended X-ray absorption fine structure (EXAFS) spectroscopy (Fig. 3j) show a dominant peak at 2.2 Å, indexed to the Y–S bonds in comparison with yttrium sulfide. Theoretically, the energy of 2H WS<sub>2</sub> is much lower than its 1T phase (>0.6 eV per formula unit). With increasing Y doping levels, the energy differences between 2H and 1T phases noticeably decrease, but it remains difficult to reverse the relative stability (Supplementary Fig. 68). Owing to the substitution of W (valency +4) in WS<sub>2</sub> by low-valence Y (+3), the charge densities near the Fermi level tend to be localized around Y atoms, thus going against electron

transfer (Supplementary Fig. 69). Moreover, the increase of defect scattering by Y dopants would also reduce the electrical conductivity. Thus, the electrical conductivity of Y-doped WS<sub>2</sub> is experimentally measured to be  $2.31 \times 10^{-3} \text{ S cm}^{-1}$ , much lower than that of pure WS<sub>2</sub> ( $8.13 \times 10^{-2} \text{ S cm}^{-1}$ ) (Supplementary Table 3). This is in contrast to those reported for *d* orbital electron-enriched metal (Re, Nb)-doped MoS<sub>2</sub><sup>31,32</sup> and our Nb-doped TiSe<sub>2</sub> (Supplementary Figs. 47–50), exhibiting substantial improvements in the electrical conductivities (Supplementary Figs. 51, 70 and Supplementary Table 3).

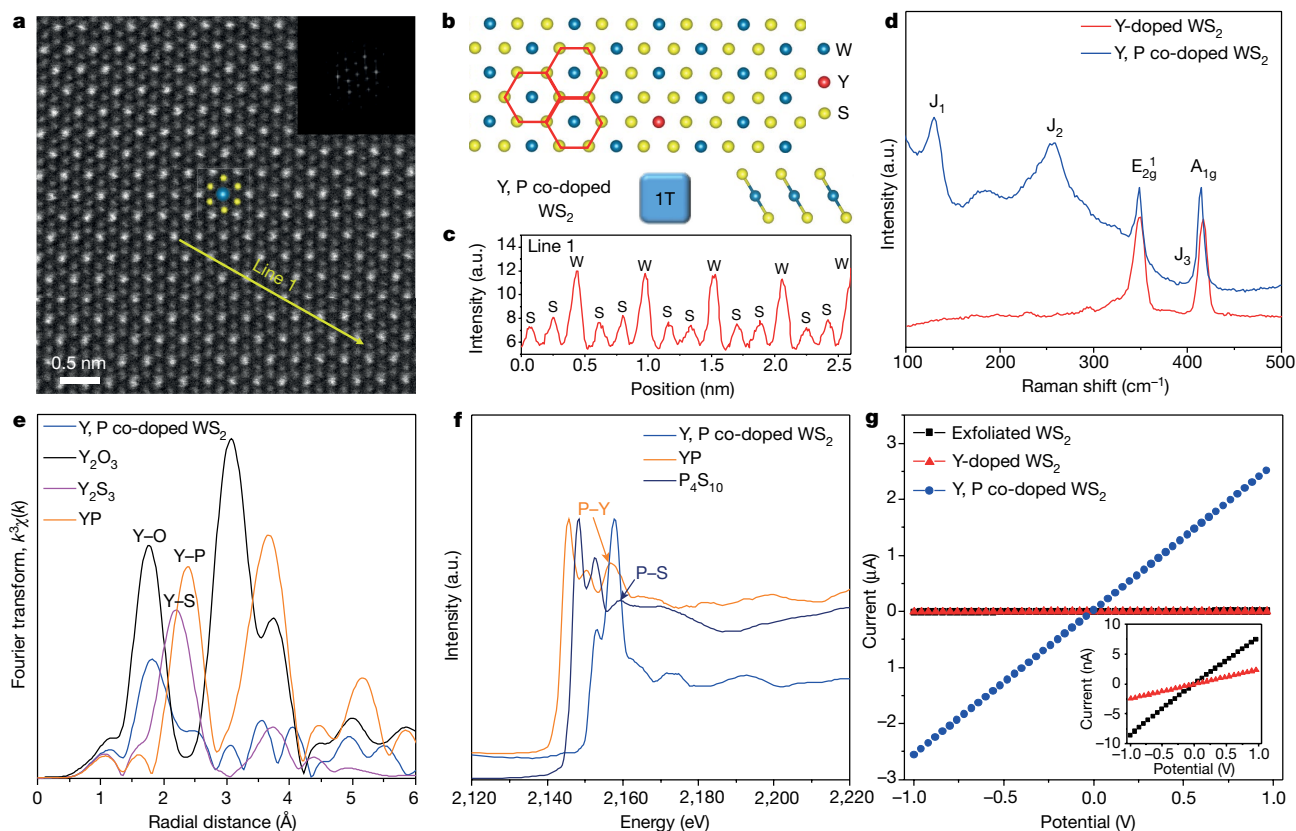
## 2D heteroatoms (Y, P) co-doped WS<sub>2</sub> (1T)

While engineering the quaternary MAX phase (W<sub>2/3</sub>Y<sub>1/3</sub>)<sub>2</sub>AlC, other vapours such as phosphorus could be easily and simultaneously introduced into the synthetic system, creating both Y and P co-doped WS<sub>2</sub> (Supplementary Figs. 52–56). As illustrated by the Raman spectra (Fig. 4d), for Y, P co-doped WS<sub>2</sub> there are three dominant peaks at 130 cm<sup>-1</sup> (J<sub>1</sub>), 258 cm<sup>-1</sup> (J<sub>2</sub>) and 406 cm<sup>-1</sup> (J<sub>3</sub>), corresponding to the vibration modes of the 1T phase<sup>6,33</sup>, as well as two peaks at 348 cm<sup>-1</sup> and 414 cm<sup>-1</sup>, indexed to the E<sub>2g</sub><sup>1</sup> and A<sub>1g</sub> peaks of the 2H phase, respectively. The presence of the 1T phase in real space can be visualized via aberration-corrected STEM images (Fig. 4a and Supplementary



**Fig. 3 | Structural characterization of 2D heteroatom-doped transition-metal chalcogenides with 2H phase derived from quaternary MAX phases.** **a**, Schematic illustration of the conversion of quaternary MAX-(W<sub>2/3</sub>Y<sub>1/3</sub>)<sub>2</sub>AlC to Y-doped WS<sub>2</sub>. **b**, Typical scanning electron microscope image of Y-doped WS<sub>2</sub>, showing the expanded structure. **c**, **d**, High-resolution TEM (**c**) and high-angle

annular dark-field (**d**) images of Y-doped WS<sub>2</sub> layers. **e–g**, Elemental mapping images of W (**e**), Y (**f**) and S (**g**) species in Y-doped WS<sub>2</sub> layers. **h**, Raman spectra of Y-doped WS<sub>2</sub> and bulk WS<sub>2</sub>. **i**, **j**, Y K-edge XANES spectra (**i**) and Fourier transform spectra (**j**) of Y K-edge EXAFS for Y-doped WS<sub>2</sub>, demonstrating the presence of Y–S bonds in Y-doped WS<sub>2</sub>.



**Fig. 4 | Structural characterization and electrical properties of 2D heteroatom (Y and P) co-doped WS<sub>2</sub> with 1T phase derived from quaternary MAX-(W<sub>2/3</sub>Y<sub>1/3</sub>)<sub>2</sub>AlC.** **a**, Atomic-resolution STEM image and its corresponding fast Fourier transform patterns (inset). **b**, Atomic configuration of Y, P co-doped WS<sub>2</sub>, exhibiting the 1T phase. **c**, Line intensity profile along the

highlighted arrow in **a**. **d**, Raman spectra of Y, P co-doped WS<sub>2</sub>. **e**, Fourier transform spectra of Y K-edge EXAFS. **f**, P K-edge XANES spectra for Y, P co-doped WS<sub>2</sub>. **g**, Current versus voltage curves of Y, P co-doped WS<sub>2</sub>, Y-doped WS<sub>2</sub> and exfoliated WS<sub>2</sub>.

Fig. 57). Combined with the intensities of cation and anion sites, the S and W elements can be identified using image contrast, as shown by the line intensity profile (Fig. 4c) acquired along the highlighted arrow in Fig. 4a. All the metal W atoms are located at the centres of octahedral units, in accordance with the atomic models of the 1T phase<sup>4,34</sup> (Fig. 4b). Y elemental mapping images (Supplementary Figs. 58a and 59a) show that all the Y atoms are located at the centres of octahedral units, occupying the W sites in WS<sub>2</sub>. This can be further demonstrated via the line intensity profiles (Supplementary Figs. 58c, 58e and 59c), in which the intensities of the Y sites are weaker than those of the W sites and stronger than those of the S sites, directly demonstrating the presence of Y–S bonds in Y, P co-doped WS<sub>2</sub>. By carefully analysing the Y, P elemental mapping images and the line intensity profiles (Supplementary Figs. 58f, 58g, 59d–i), it is clear that the P atoms are located on the top of both Y and S sites, verifying the presence of Y–P and P–S bonds in the sample. These bonds can be further demonstrated via Fourier transform spectra of EXAFS (Fig. 4e and Supplementary Fig. 60), where there are two dominant peaks between 1.3 Å and 2.8 Å, corresponding to the overlap of Y–O, Y–S and Y–P bonds at 1.7 Å, 2.2 Å and 2.4 Å, respectively. In the P K-edge XANES spectra of Y, P co-doped WS<sub>2</sub> (Fig. 4f), there is one prominent peak centred at 2,158 eV, attributed to the overlap of P–Y bonds (2,157 eV) and P–S bonds (2,159 eV), in good agreement with the above EXAFS, aberration-corrected STEM images and the corresponding elemental mapping analysis.

Density functional theory calculations confirm that as P atoms adsorb on the hollow site 2 (h2), sulfur site 2 (S2) and tungsten site (W), the energy differences between 2H and 1T WS<sub>2</sub> are negligible. However, when P atoms adsorb on the top of the Y atoms and their neighbours such as hollow site 1 (h1) and sulfur site 1 (S1), the energies of the 1T WS<sub>2</sub> are surprisingly lower than when in the 2H phase, revealing that there is a unique yttrium–phosphorus (Y–P) joint effect that stabilizes the configuration of the 1T phase (Supplementary Figs. 71 and 72). Furthermore, it is difficult to reverse the relative energy stability between 2H and 1T WS<sub>2</sub> by independent Y-doping or P-adsorption (Supplementary Figs. 61 and 73). Even after storage under ambient conditions for about one year, the 1T-containing WS<sub>2</sub> remains stable (Supplementary Fig. 62), unlike the 1T transition-metal dichalcogenides produced via traditional methods that have poor stability at high temperatures (>573 K)<sup>8,35</sup>. Such Y, P co-doped WS<sub>2</sub> exhibits a linear current–voltage (*I*–*V*) characteristic with a low resistance of 387 kΩ per □, close to that reported for 1T' WS<sub>2</sub> (430 kΩ per □)<sup>6</sup>, three orders of magnitude lower than those of 2H Y-doped WS<sub>2</sub> (413 MΩ per □) and exfoliated WS<sub>2</sub> (124 MΩ per □) (Fig. 4g, Supplementary Fig. 63 and Supplementary Table 3).

The accordion-like structure, highly exposed surfaces and abundant 1T phase (Fig. 4d and Supplementary Fig. 64) of the resultant transition-metal dichalcogenides mean that they could be directly used as electrocatalysts for the hydrogen evolution reaction (Supplementary Figs. 65 and 66). We believe that our synthetic protocol has the potential to convert a series of non-vdW solids to 2D vdW nanocrystals with selected phases, achieving high-throughput monolayers, specific dopants and tailored electronic features, as well as broad applications in fields such as electronics, catalysis and energy storage.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-019-1904-x>.

- Coleman, J. N. et al. Two-dimensional nanosheets produced by liquid exfoliation of layered materials. *Science* **331**, 568–571 (2011).
- Lin, Z. et al. Solution-processable 2D semiconductors for high-performance large-area electronics. *Nature* **562**, 254–258 (2018).
- Huang, Y. et al. Reliable exfoliation of large-area high-quality flakes of graphene and other two-dimensional materials. *ACS Nano* **9**, 10612–10620 (2015).
- Zhou, J. et al. A library of atomically thin metal chalcogenides. *Nature* **556**, 355–359 (2018).
- Chen, X. et al. CVD-grown monolayer MoS<sub>2</sub> in bioabsorbable electronics and biosensors. *Nat. Commun.* **9**, 1690–1701 (2018).
- Voiry, D. et al. Enhanced catalytic activity in strained chemically exfoliated WS<sub>2</sub> nanosheets for hydrogen evolution. *Nat. Mater.* **12**, 850–855 (2013).
- Zhu, J. et al. Argon plasma induced phase transition in monolayer MoS<sub>2</sub>. *J. Am. Chem. Soc.* **139**, 10216–10219 (2017).
- Eda, G. et al. Photoluminescence from chemically exfoliated MoS<sub>2</sub>. *Nano Lett.* **11**, 5111–5116 (2011).
- Zhang, J. et al. Single platinum atoms immobilized on an MXene as an efficient catalyst for the hydrogen evolution reaction. *Nat. Catal.* **1**, 985–992 (2018).
- Kundu, D., Adams, B. D., Ort, V. D., Vajargah, S. H. & Nazar, L. F. A high-capacity and long-life aqueous rechargeable zinc battery using a metal oxide intercalation cathode. *Nat. Energy* **1**, 16119–16126 (2016).
- Zhang, C. et al. High capacity silicon anodes enabled by MXene viscous aqueous ink. *Nat. Commun.* **10**, 849–857 (2019).
- Yan, M. et al. Water-lubricated intercalation in V<sub>2</sub>O<sub>5</sub>·nH<sub>2</sub>O for high-capacity and high-rate aqueous rechargeable zinc batteries. *Adv. Mater.* **30**, 1703725–1703730 (2018).
- Wang, Y. et al. In situ exfoliated, N-doped, and edge-rich ultrathin layered double hydroxides nanosheets for oxygen evolution reaction. *Adv. Funct. Mater.* **28**, 1703363–1703368 (2018).
- Liu, J., Qian, X. & Fu, L. Crystal field effect induced topological crystalline insulators in monolayer IV–VI semiconductors. *Nano Lett.* **15**, 2657–2661 (2015).
- Balan, A. P. et al. Exfoliation of a non-van der Waals material from iron ore hematite. *Nat. Nanotechnol.* **13**, 602–609 (2018).
- Naguib, M. et al. Two-dimensional nanocrystals produced by exfoliation of Ti<sub>3</sub>AlC<sub>2</sub>. *Adv. Mater.* **23**, 4248–4253 (2011).
- Barsoum, M. W. *MAX Phases: Properties of Machinable Ternary Carbides and Nitrides* (Wiley, 2013).
- Sokol, M., Natu, V., Kota, S. & Barsoum, M. W. On the chemical diversity of the MAX phases. *Trends Chem.* **1**, 210–223 (2019).
- Xuan, J. et al. Organic-base-driven intercalation and delamination for the production of functionalized titanium carbide nanosheets with superior photothermal therapeutic performance. *Angew. Chem. Int. Ed.* **55**, 14569–14574 (2016).
- Ghidiu, M., Lukatskaya, M. R., Zhao, M. Q., Gogotsi, Y. & Barsoum, M. W. Conductive two-dimensional titanium carbide ‘clay’ with high volumetric capacitance. *Nature* **516**, 78–81 (2014).
- Lukatskaya, M. R. et al. Cation intercalation and high volumetric capacitance of two-dimensional titanium carbide. *Science* **341**, 1502–1505 (2013).
- Li, M. et al. Element replacement approach by reaction with Lewis acidic molten salts to synthesize nanolaminated MAX phases and MXenes. *J. Am. Chem. Soc.* **141**, 4730–4737 (2019).
- Barin, I. *Thermochemical Data of Pure Substances* 3rd edn (VCH, 1995).
- Yaws, C. L. *The Yaws Handbook of Physical Properties for Hydrocarbons and Chemicals* 2nd edn (Elsevier, 2015).
- Winterbone, D. E. *Advanced Thermodynamics for Engineers* (Butterworth-Heinemann, 1997).
- Naguib, M., Mochalin, V. N., Barsoum, M. W. & Gogotsi, Y. 25th anniversary article: MXenes: a new family of two-dimensional materials. *Adv. Mater.* **26**, 992–1005 (2014).
- Anasori, B., Lukatskaya, M. R. & Gogotsi, Y. 2D metal carbides and nitrides (MXenes) for energy storage. *Nat. Rev. Mater.* **2**, 16098–16114 (2017).
- Anasori, B. et al. Two-dimensional, ordered, double transition metals carbides (MXenes). *ACS Nano* **9**, 9507–9516 (2015).
- Meshkian, R. et al. W-based atomic laminates and their 2D derivative W<sub>133</sub>C MXene with vacancy ordering. *Adv. Mater.* **30**, 1706409–1706416 (2018).
- Naguib, M. et al. Two-dimensional transition metal carbides. *ACS Nano* **6**, 1322–1331 (2012).
- Deepak, F. L. et al. Fullerene-like (IF) Nb<sub>x</sub>Mo<sub>1-x</sub>S<sub>2</sub> nanoparticles. *J. Am. Chem. Soc.* **129**, 12549–12562 (2007).
- Yang, S. Z. et al. Rhenium-doped and stabilized MoS<sub>2</sub> atomic layers with basal-plane catalytic activity. *Adv. Mater.* **30**, 1803477–1803483 (2018).
- Liu, Q. et al. Stable metallic 1T-WS<sub>2</sub> nanoribbons intercalated with ammonia ions: the correlation between structure and electrical/optical properties. *Adv. Mater.* **27**, 4837–4844 (2015).
- Radhakrishnan, S. et al. An insight into the phase transformation of WS<sub>2</sub> upon fluorination. *Adv. Mater.* **30**, 1803366–1803375 (2018).
- Lin, Y. C., Dumcenco, D. O., Huang, Y. S. & Suenaga, K. Atomic mechanism of the semiconducting-to-metallic phase transition in single-layered MoS<sub>2</sub>. *Nat. Nanotechnol.* **9**, 391–396 (2014).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2020



## Methods

### Synthesis of MAX phases

Some MAX phases were synthesized by ball-milling of commercially available powders and subsequent calcination treatments<sup>36–40</sup>. Taking Mo<sub>2</sub>GeC as an example, commercial Mo, Ge and graphite in a molar ratio of 2:1.05:1 were sealed in an agate container with agate balls and milled at 600 rpm for 20 h. The mixture was then heated at a rate of 3 K min<sup>−1</sup> until it reached 1,673 K and was maintained at this high temperature for 4 h. After cooling to room temperature, the bulk was ground to produce MAX-Mo<sub>2</sub>GeC. For other MAX phases, the details are listed in Materials and Methods in the Supplementary Information.

### Synthesis of 2D transition-metal chalcogenides

Transition-metal sulfides were prepared by the reaction of MAX phases or MoB with H<sub>2</sub>S gas at temperatures of 1,073–1,373 K. Taking a TMD-MoS<sub>2</sub> as an example, 300 mg of MAX-Mo<sub>2</sub>GeC was heated at a heating rate of 10 K min<sup>−1</sup> under Ar flow, and an H<sub>2</sub>S/Ar (10 vol.% H<sub>2</sub>S) mixture was injected when the temperature reached 1,073 K. MAX-Mo<sub>2</sub>GeC was maintained at this temperature for 4 h, generating TMD-MoS<sub>2</sub>. Transition metal selenides were prepared by the reaction of MAX phases, MoB and MoSi<sub>2</sub> with Se vapours at temperatures of 1,073–1,373 K. Specifically, 2 g of Se powder and 300 mg of MAX-Ti<sub>3</sub>SiC<sub>2</sub> were placed in low- (973 K) and high- (1,173–1,273 K) temperature zones, respectively.

### Synthesis of 2D heteroatom-doped transition-metal chalcogenides

Heteroatom-doped transition-metal chalcogenides were synthesized by the reaction of quaternary MAX phases with chalcogen-containing gases. Specifically, 300 mg of MAX-(W<sub>2/3</sub>Y<sub>1/3</sub>)<sub>2</sub>AlC was heated at a heating rate of 10 K min<sup>−1</sup> under Ar flow, and then H<sub>2</sub>S/Ar mixture was injected when the temperature reached 1,273 K. MAX-(W<sub>2/3</sub>Y<sub>1/3</sub>)<sub>2</sub>AlC was maintained there for 4 h to produce Y-doped WS<sub>2</sub>. Nb-doped TiSe<sub>2</sub> was derived from MAX-(Ti<sub>1/2</sub>Nb<sub>1/2</sub>)<sub>2</sub>AlC using the same procedures as for making TMD-TiSe<sub>2</sub>.

### Synthesis of 2D heteroatoms (Y and P) co-doped WS<sub>2</sub>

Y, P co-doped WS<sub>2</sub> was synthesized by the reaction of MAX-(W<sub>2/3</sub>Y<sub>1/3</sub>)<sub>2</sub>AlC with H<sub>2</sub>S gas and P vapour at a high temperature of 1,273 K. Specifically, 300 mg of (W<sub>2/3</sub>Y<sub>1/3</sub>)<sub>2</sub>AlC and 1 g of P were placed into two separate crucibles, where P powder was placed in a low-temperature upstream zone maintained at 873 K. (W<sub>2/3</sub>Y<sub>1/3</sub>)<sub>2</sub>AlC was heated at a heating rate of 10 K min<sup>−1</sup> under Ar flow, and then the H<sub>2</sub>S/Ar mixture was injected when the temperature reached 1,273 K. (W<sub>2/3</sub>Y<sub>1/3</sub>)<sub>2</sub>AlC was maintained at 1,273 K for 4 h to produce Y, P co-doped WS<sub>2</sub>. Similarly, P-doped MoS<sub>2</sub> was prepared by the reaction of MAX-Mo<sub>2</sub>GeC with H<sub>2</sub>S gas and P vapour at a high temperature of 1,073 K.

### Synthesis of P-doped WS<sub>2</sub>

P-doped WS<sub>2</sub> was synthesized by the reaction of bulk WS<sub>2</sub> with P vapour at high temperature. Specifically, 300 mg of WS<sub>2</sub> powders were put in a porcelain boat with 1 g of P at the upstream zone. Then the boat was heated to 1,273 K at a heating rate of 10 K min<sup>−1</sup> under Ar flow and kept there for 30 min to generate P-doped WS<sub>2</sub>.

### Fabrication of thin films of Y, P co-doped WS<sub>2</sub>, Y-doped WS<sub>2</sub> and exfoliated WS<sub>2</sub>

Y, P co-doped WS<sub>2</sub> thin film was fabricated by vacuum filtration of Y, P co-doped WS<sub>2</sub> nanosheets on nylon membrane filters, which were acquired by a facile liquid exfoliation of accordion-like Y, P co-doped WS<sub>2</sub> in an isopropyl alcohol solvent<sup>1</sup>. Other thin films of Y-doped WS<sub>2</sub>, WS<sub>2</sub>, Nb-doped TiSe<sub>2</sub> and TiSe<sub>2</sub> were similarly obtained.

### Characterization

The morphology and microstructure of materials were characterized by scanning electron microscopy (Zeiss MERLIN Compact), transmission

electron microscopy (JEOL 2100F), spherical aberration-corrected transmission electron microscopy (FEI Titan G2) and X-ray diffraction (Rigaku D/MAX2200pc). Raman spectra were recorded on a Renishaw inVia Microscopic confocal Raman spectrometer using a 532-nm laser beam. X-ray photoelectron spectroscopy was recorded by a Thermo Electron ESCALAB 250 XPS spectrometer. Atomic force microscopy measurements were carried out on a Dimension ICON scanning probe microscope (Veeco/Bruker). X-ray absorption near-edge fine structure (XANES) and extended X-ray absorption fine structure (EXAFS) data for the Y K-edge were collected on BL14W1 and BL1W1B at the Shanghai Synchrotron Radiation Facility and the Beijing Synchrotron Radiation Facility, respectively. XANES data for the P K-edge were collected on BL4B7B at the Beijing Synchrotron Radiation Facility. Current-versus-voltage measurements were conducted using the two-electrode method on an electrochemical workstation (CHI760E, CH Instruments) in a voltage range of −1 V to 1 V at a scan rate of 10 mV s<sup>−1</sup>. The electrical conductivities of powder samples were investigated on a four-probe powder resistivity tester (ST2722-SZ, Suzhou Jingge Electronic Co., Ltd).

### Vapour pressure calculations

With the aid of log  $K_f$  values for two-phase equilibria solid–gas or liquid–gas ( $AZ_{\text{sol,liq}}-AZ_{\text{gas}}$ ), we calculated the vapour pressures of AZ gases according to the following equation<sup>23</sup>:

$$\log K = \log K_f(AZ_{\text{gas}}) - \log K_f(AZ_{\text{sol,liq}}) = \log[p(AZ_{\text{gas}})/a(AZ_{\text{sol,liq}})]$$

where  $p$  is the vapour pressure,  $a$  is the activity of AZ in the condensed phase,  $K_f$  is the equilibrium constant of formation reaction and log  $K_f$  values taken from the literature<sup>23</sup> are partially listed in Supplementary Tables 4–7. In general,  $a = 1$  for pure substances in a condensed phase.

For the evaporation equation of pure substances at a given temperature  $T$ :

$$\log K_f(AZ_{\text{gas}}) - \log K_f(AZ_{\text{sol,liq}}) = \log p(AZ_{\text{gas}})$$

and

$$p(AZ_{\text{gas}}) = 10^{\log K_f(AZ_{\text{gas}}) - \log K_f(AZ_{\text{sol,liq}})}$$

where  $p$  is in units of bar (1 bar = 10<sup>5</sup> Pa). Taking  $1/T$  and  $\ln p$  as the horizontal and vertical axis respectively, linear curves were made as shown in Supplementary Fig. 6. According to the Clausius–Clapeyron equation<sup>25</sup>:

$$\ln p = -\Delta H_m/(RT) + C \quad (1)$$

where  $p$  is vapour pressure,  $\Delta H_m$  is molar enthalpy of evaporation,  $R$  is molar gas constant and  $C$  is a constant, these linear curves confirm that the Clausius–Clapeyron equation can be applied to evaluate the relationship between vapour pressure and temperature.

In the case of AZ substances that lack log  $K_f$  values in the literature, the boiling point and the corresponding vapour pressure (1 atm = 101,325 Pa) are selected to evaluate the relationship between vapour pressure and temperature (Supplementary Table 8)<sup>24</sup>.

### Data availability

The data that support the findings of this study are available from the corresponding authors on reasonable request.

36. Sun, Z. M. Progress in research and development on MAX phases: a family of layered ternary compounds. *Int. Mater. Rev.* **56**, 143–166 (2011).

37. Barsoum, M. W. The  $M_{n+1}AX_n$  phases: a new class of solids: thermodynamically stable nanolaminates. *Prog. Solid State Chem.* **28**, 201–281 (2000).

38. Eklund, P., Beckers, M., Jansson, U., Hogberg, H. & Hultman, L. The  $M_{n+1}AX_n$  phases: materials science and thin-film processing. *Thin Solid Films* **518**, 1851–1878 (2010).



39. Magnuson, M. & Mattesini, M. Chemical bonding and electronic-structure in MAX phases as viewed by X-ray spectroscopy and density functional theory. *Thin Solid Films* **621**, 108–130 (2017).
40. Naguib, M. et al. New two-dimensional niobium and vanadium carbides as promising materials for Li-ion batteries. *J. Am. Chem. Soc.* **135**, 15966–15969 (2013).

**Acknowledgements** This work was financially supported by the National Natural Science Foundation of China (grant numbers 51622203 and 51572007), the Youth 1000-Talent Program of China and the 111 Project (grant number B17002). X.Z. also thanks Shenzhen Basic Research Projects (grant number JCYJ20170407155608882), the China Postdoctoral Science Foundation (grant number 2018M631458) and the Development and Reform Commission of Shenzhen Municipality for the development of the “Low-Dimensional Materials and Devices” Discipline, Guangdong Innovative and Entrepreneurial Research Team Program (grant number 2017ZT07C341). We thank the Shanghai Synchrotron Radiation Facility and Beijing Synchrotron Radiation Facility for support. We thank X. Chen, Q. Zhang, W. Zhou and M. Li for help with the TEM analysis; S. Chen and Y. Lin for help with the EXAFS analysis; and L. Ma for suggestions.

**Author contributions** S.Y., S.L. and P.M.A. supervised the project. S.Y. and Z.D. designed and carried out all of the experiments. S.Z. and X.Z. carried out the density functional theory calculations. S.W. and B.L. performed the scanning electron microscope, TEM and X-ray diffraction measurements. B.L. carried out the X-ray photoelectron spectroscopy and Raman analysis. L.S. contributed to the XANES and EXAFS measurements. Z.D. performed the atomic force microscopy. Z.D. and Y.G. designed the film electrodes and carried out the electrical and electrocatalytic measurements. All authors discussed the results and assisted during manuscript preparation.

**Competing interests** The authors declare no competing interests.

#### Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41586-019-1904-x>.

**Correspondence and requests for materials** should be addressed to S.Y. or P.M.A.

**Peer review information** *Nature* thanks Per Eklund, Wei Sun Leong and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>.

# Antagonistic cooperativity between crystal growth modifiers

<https://doi.org/10.1038/s41586-019-1918-4>

Wenchuan Ma<sup>1</sup>, James F. Lutsko<sup>2</sup>, Jeffrey D. Rimer<sup>1,3\*</sup> & Peter G. Vekilov<sup>1,3\*</sup>

Received: 11 February 2019

Accepted: 17 October 2019

Published online: 15 January 2020

Ubiquitous processes in nature and the industry exploit crystallization from multicomponent environments<sup>1–5</sup>; however, laboratory efforts have focused on the crystallization of pure solutes<sup>6,7</sup> and the effects of single growth modifiers<sup>8,9</sup>. Here we examine the molecular mechanisms employed by pairs of inhibitors in blocking the crystallization of haematin, which is a model organic compound with relevance to the physiology of malaria parasites<sup>10,11</sup>. We use a combination of scanning probe microscopy and molecular modelling to demonstrate that inhibitor pairs, whose constituents adopt distinct mechanisms of haematin growth inhibition, kink blocking and step pinning<sup>12,13</sup>, exhibit both synergistic and antagonistic cooperativity depending on the inhibitor combination and applied concentrations. Synergism between two crystal growth modifiers is expected, but the antagonistic cooperativity of haematin inhibitors is not reflected in current crystal growth models. We demonstrate that kink blockers reduce the line tension of step edges, which facilitates both the nucleation of crystal layers and step propagation through the gates created by step pinners. The molecular viewpoint on cooperativity between crystallization modifiers provides guidance on the pairing of modifiers in the synthesis of crystalline materials. The proposed mechanisms indicate strategies to understand and control crystallization in both natural and engineered systems, which occurs in complex multicomponent media<sup>1–3,8,9</sup>. In a broader context, our results highlight the complexity of crystal–modifier interactions mediated by the structure and dynamics of the crystal interface.

Crystallization is the central process of materials synthesis in biological, geological and extraterrestrial systems<sup>7,14</sup>. Nature achieves a remarkable diversity of shapes, patterns, compositions and functions of the arising crystalline structures by combining simple strategies to control the number of nucleated crystals and their anisotropic rates of growth<sup>15,16</sup>. To promote or inhibit crystallization in both natural and engineered environments, soluble foreign compounds are deployed that interact with the solute or the crystal/solution interface<sup>17</sup>. In many cases, two or more modifiers operate in tandem to alter the processes of crystallization<sup>4,18–21</sup>; however, the fundamental mode(s) of cooperative action is not well understood.

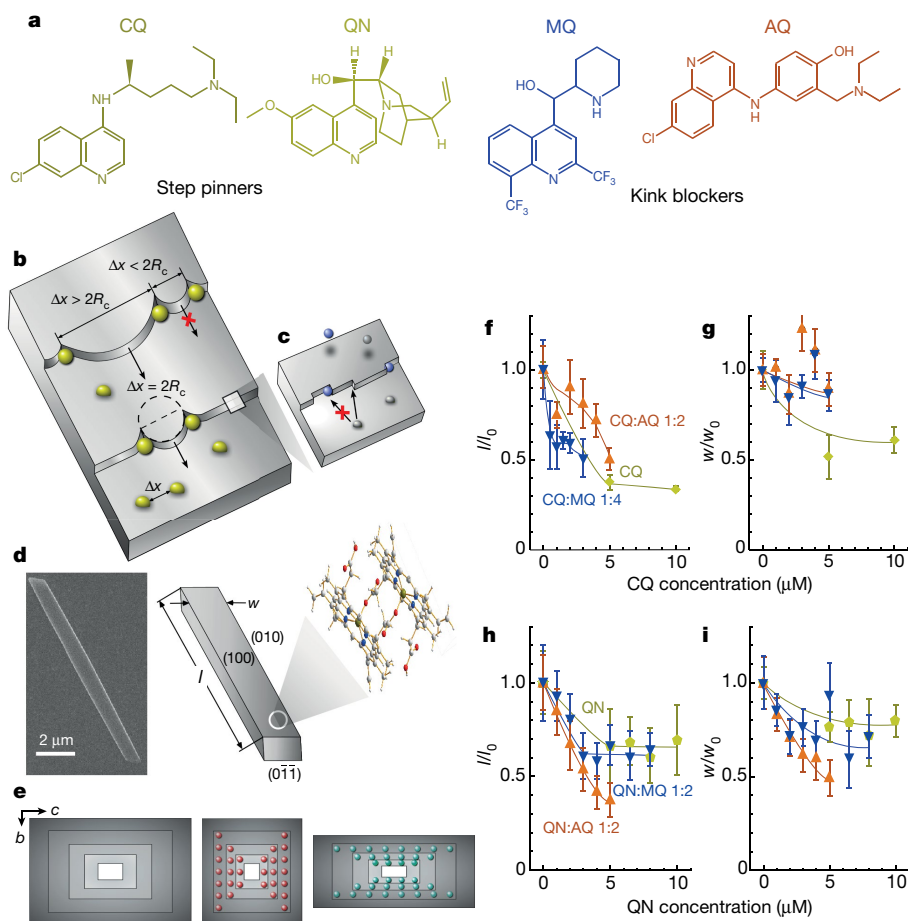
To gain molecular-level insight into the mechanisms of cooperativity between crystallization modifiers, we examine the growth of  $\beta$ -haematin crystals, which form in malaria parasites as a part of their haem-detoxification mechanism<sup>22</sup>, in the presence of quinoline compounds, which represent a major class of the currently employed antimalarials<sup>23,24</sup>. Recent work has established that  $\beta$ -haematin crystal growth follows classical mechanisms, whereby new layers nucleate on the crystal surfaces and advance by incorporation of solute molecules at the steps<sup>12</sup>. These studies uncovered two distinct classes of quinoline inhibition of step propagation<sup>13</sup>. In the first mechanism, known as ‘step pinning’, chloroquine (CQ) and quinine (QN; Fig. 1a) bind to flat terraces

and arrest crystal formation over broad areas of the crystal surface (Fig. 1b)<sup>25</sup>. In addition, amodiaquine (AQ) and mefloquine (MQ; Fig. 1a) were found to block kinks, the sites where haematin molecules incorporate into steps (Fig. 1c)<sup>12</sup>.

Even though combinations of two or more crystal growth inhibitors are common in many drug formulations<sup>26</sup>, a crucial gap in the understanding of interactions between inhibitor pairs that regulate haematin crystallization has been identified<sup>27,28</sup>. To address the molecular mechanism of action of binary inhibitor combinations on  $\beta$ -haematin crystal growth, we pair a step pinner, CQ or QN, with a kink blocker, MQ or AQ. We classify the cooperativity between paired inhibitors as synergistic, additive or antagonistic according to whether the response to a combination of two inhibitors is, respectively, stronger, equal or weaker than the sum of the responses to individual doses<sup>29</sup>.

Binary inhibitor combinations impose dramatic changes in the shapes and dimensions of  $\beta$ -haematin crystals (Fig. 1f–i, Extended Data Fig. 1). The crystal length along the *c* crystallographic axis is the result of growth in the [011] and [0 $\bar{1}$ 1] directions (Fig. 1d, e). The shorter average length enforced by both MQ and CQ than by either modifier separately indicates a strong synergistic activity of these two inhibitors (Fig. 1f). As the crystal length is insensitive to the presence of MQ alone<sup>13</sup>,

<sup>1</sup>Department of Chemical and Biomolecular Engineering, University of Houston, Houston, TX, USA. <sup>2</sup>Center for Nonlinear Phenomena and Complex Systems, Université Libre de Bruxelles, Brussels, Belgium. <sup>3</sup>Department of Chemistry, University of Houston, Houston, TX, USA. \*e-mail: jrimer@uh.edu; vekilov@uh.edu



**Fig. 1 | Cooperativity between four pairs of inhibitors in suppressing bulk growth of  $\beta$ -haematin crystals.** **a**, Structures of step pinners CQ and QN and kink blockers MQ and AQ. **b**, Schematic of step pinning, where  $\Delta x$  is the separation between inhibitor molecules (shown in gold) adsorbed on flat crystal terraces and  $R_c$  is the critical radius of the two-dimensional nucleus. Step growth is delayed if  $\Delta x$  is comparable to  $2R_c$  and arrested if  $\Delta x < 2R_c$ . **c**, Schematic of inhibitors (shown in blue) inhibiting step advancement by partial blocking of access of solute molecules to kinks. **d**, Scanning electron microscopy micrograph and schematic illustrating the  $\beta$ -haematin crystal shape.  $l$ , length;  $w$ , width. **e**, Preservation of the crystal shape during growth in pure solutions and inhibitor-induced suppression of crystal length or width by

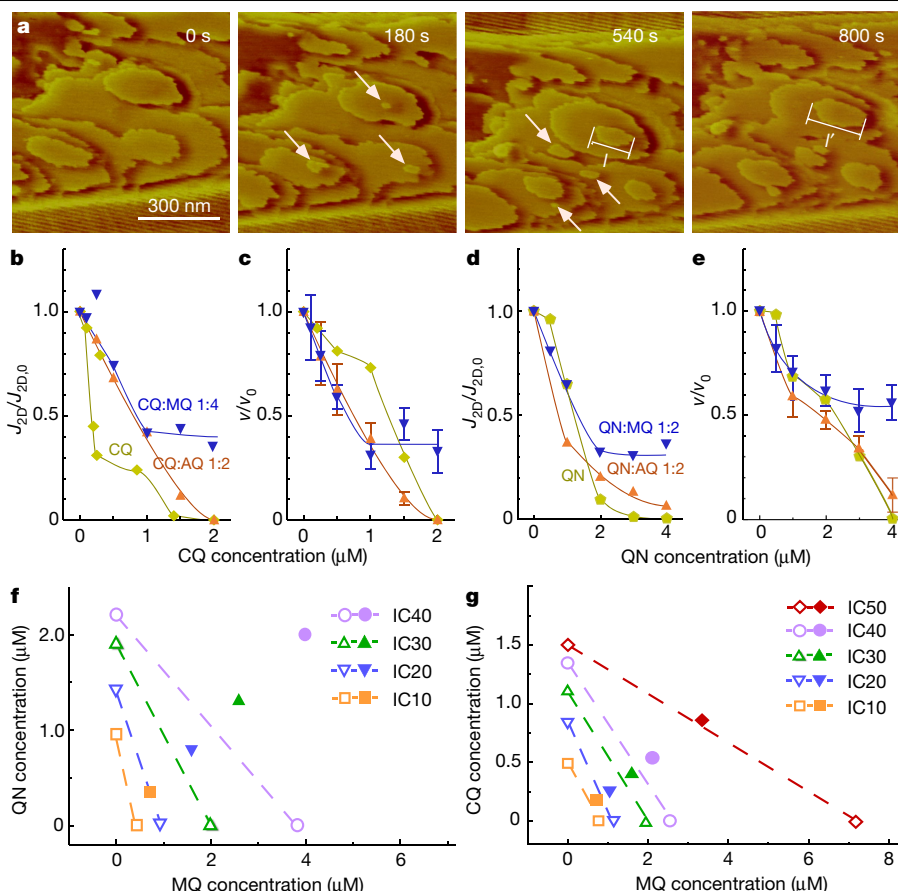
additive cooperativity of CQ and MQ engenders crystal lengths similar to those constrained by only CQ. By contrast, the crystal lengths affected by the pairing of AQ and CQ are substantially longer than those engendered by only CQ, implying an antagonistic cooperativity between these two modifiers. The addition of either MQ or AQ to CQ-containing solutions enforces greater crystal widths than those with only CQ (Fig. 1g). The crystal width increases owing to growth in the  $\langle 010 \rangle$  directions (Fig. 1d); thus, greater widths indicate that the MQ/CQ and AQ/CQ pairs impede growth of  $\{010\}$  faces to a lesser extent than CQ on its own. We previously reported that MQ and AQ weakly affect the crystal width<sup>13</sup>; therefore, these new findings indicate antagonistic cooperativity of CQ with kink blockers MQ and AQ in inhibiting the width of  $\beta$ -haematin crystals. Notably, in select inhibitor concentration ranges (for example,  $C_{CQ} < 1 \mu\text{M}$  and  $C_{MQ} < 4 \mu\text{M}$ ) synergism in suppressing growth along the  $c$  axis accompanies antagonistic cooperativity towards growth in the  $b$  direction (Fig. 1f, g); the opposite responses are probably defined by the selective binding of the inhibitors to the individual crystal faces dictated by their distinct structures<sup>17,21</sup>. Importantly, they further weaken the synergistic cooperativity of CQ and MQ in inhibiting haematin sequestration into

interaction of inhibitors with axial and lateral crystal faces, respectively. Red spheres denote inhibitors of the  $c$  face; green spheres denote inhibitors of the  $b$  face. **f–i**, Variations of the average length  $l$  and width  $w$  of crystals grown in the presence of increasing concentrations of four inhibitor pairs at the displayed ratios relative to  $l_0$  and  $w_0$ , reached after growth in pure growth solutions for 16 d at 23 °C. Error bars represent the standard deviation of about 30 measurements. Lines are guides for the eye. In all experiments, haematin concentration  $c_H = 0.28 \text{ mM}$  and supersaturation  $\sigma = \ln(c_H/c_e) \approx 0.93$ , where  $c_e = 0.11 \text{ mM}$  is the solubility at 23 °C. The majority of the length and width data for individual modifiers are from Olafson et al.<sup>13</sup> and are consistent with additional measurements of the effects of QN.

crystals. Combining MQ and AQ with QN elicits mostly synergistic responses of both the crystal length (Fig. 1h) and width (Fig. 1i).

Antagonistic cooperativity between crystallization inhibitors appears counterintuitive. To understand the effects of inhibitor combinations on the molecular processes of growth of the  $\{100\}$  face of  $\beta$ -haematin crystals, we used time-resolved in situ atomic force microscopy (AFM)<sup>12,13</sup>. We scrutinized inhibitor effects on the rate of two-dimensional nucleation of new crystal layers  $J_{2D}$  and the rate of propagation of steps  $v$ . For  $J_{2D}$ , we counted the number of new layer nuclei that grow above a critical radius  $R_c$  per unit area of the surface and unit time (Fig. 2a). We determined  $v$  from the displacement of the steps over time (Fig. 2a)<sup>12</sup>. The correlation between  $J_{2D}$  and the concentration of the inhibitors demonstrates that the addition of the kink blockers MQ and AQ to the step pinner CQ substantially enhances the nucleation of new layers relative to that with solitary CQ, indicating strong antagonism (Fig. 2b). The cooperativity between CQ/MQ and CQ/AQ in suppressing  $v$  is antagonistic at almost all tested inhibitor concentrations (Fig. 2c, Extended Data Fig. 2b, Extended Data Table 1).

MQ and AQ exhibit a similar transition towards stronger antagonism when combined with QN. MQ, which alone does not suppress  $J_{2D}$  (ref. <sup>13</sup>),



**Fig. 2 | Cooperativity of inhibitor pairs in suppressing layer generation and spreading.** **a**, Time-resolved in situ AFM images showing the nucleation and growth of new layers on a (100) face at  $c_H = 0.28$  mM and supersaturation  $\sigma = \ln(c_H/c_e) \approx 0.56$ , where  $c_e = 0.16$  mM is the solubility at 28 °C, the temperature in the AFM liquid cell. Arrows indicate newly nucleated islands that are counted to determine the rate of two-dimensional nucleation,  $J_{2D}$ . The growth of the island dimension  $l$  underlies the determination of the step velocity,  $v$ . The bright lines with striations at the top and bottom of some of the panels correspond to the crystal edges. **b–e**, Decrease in  $J_{2D}$  relative to that in the absence of any inhibitor,  $J_{2D,0}$  (**b**, **d**) and of  $v$  relative to that in the absence of any inhibitor,  $v_0$  (**c**, **e**) with increasing concentrations of CQ (**b**, **c**) and QN (**d**, **e**) inhibitor pairs at the displayed ratios. Error bars represent the standard

deviation from the average of 15 to 25 measurements of  $J_{2D}$  and 25 to 35 measurements of  $v$ , and are, in some cases, smaller than the symbol size. Lines are guides for the eye. Data for individual modifiers are from Olafson et al.<sup>13</sup>. **f**, **g**, Isobolograms characterizing the inhibition of  $v$  by QN/MQ (**f**) and CQ/MQ (**g**). Open symbols indicate the concentrations of individual inhibitors that elicit a certain percentage of inhibition, referred to as inhibitory concentrations (ICs). Dashed lines correspond to additive cooperativity between the paired inhibitors for a certain percentage of inhibition (10–50%). Solid symbols represent the concentrations of the paired inhibitors that evoke the same inhibition. Rightward shifts of the solid symbols from the respective dashed lines indicate antagonistic cooperativity. The corresponding combination index values are listed in Extended Data Table 1.

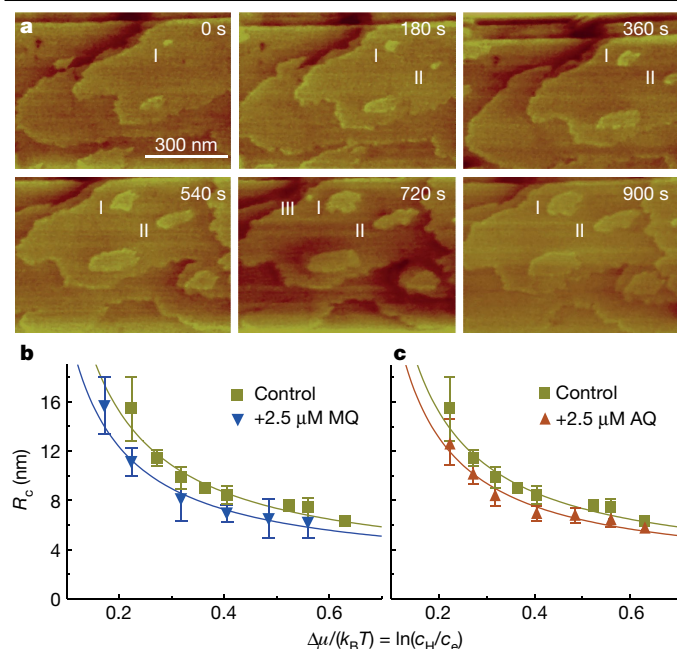
exhibits synergistic cooperativity with QN at  $C_{MQ} < 4$   $\mu$ M and antagonism at  $C_{MQ} > 4$   $\mu$ M (Fig. 2d). Similarly, AQ, which on its own depresses  $J_{2D}$  by up to 60% (ref.<sup>13</sup>), transitions from synergy at  $C_{AQ} < 2$   $\mu$ M to antagonism at  $C_{AQ} > 4$   $\mu$ M. Both MQ and AQ strongly inhibit the step velocity  $v$  when acting alone<sup>13</sup> and the similarity between velocity profiles measured in the presence of QN/MQ and QN/AQ combinations to those obstructed by only QN (Fig. 2e) signify strong antagonism between MQ and QN and between AQ and QN. The cooperativity of inhibitor pairings can be quantified from isobolograms (Fig. 2f, g), an established method in pharmaceutical research, in which the doses of paired inhibitors needed to inhibit  $J_{2D}$  and  $v$  by a certain percentage are compared with the sum of the responses to each inhibitor applied individually<sup>27,29</sup>.

We establish that the antagonism between step pinners and kink blockers in inhibiting bulk crystallization and the surface processes on (100) faces is not motivated by the formation of inhibitor–haematin complexes in solution. We examined whether the constituents of an inhibitor pair formed binary complexes that do not impede crystallization. Such complexation would lower the concentration of the active inhibitor and constrain their potency. We tested the formation of CQ/MQ, CQ/AQ, QN/MQ and QN/AQ binary complexes. Considering that

the four inhibitors form complexes with haematin<sup>13,30</sup>, we also explored whether these four combinations assemble into ternary compounds that include haematin. The results presented in Extended Data Fig. 3 show that no complexes involving both inhibitors exist in the solution, and imply that complexation between the applied inhibitors is not the source of the observed antagonistic cooperativity.

Additive and synergistic cooperativity in suppressing  $J_{2D}$  and  $v$  between a kink blocker and a step pinner can be understood within the realm of common crystal growth models. Blocking of kinks lowers the kinetic constant for growth, which works in parallel with the depression of the crystallization driving force due to step curvature enforced by step pinners (Fig. 1a, b). We posit that the antagonism between the two types of inhibitors originates from the reduction of the step line tension  $\gamma$ , a thermodynamic prerequisite for the adsorption of kink blockers at steps<sup>31</sup>. On the basis of the Gibbs–Thomson relation,  $\gamma$  regulates the radius of the critical two-dimensional nucleus according to  $R_c = \Omega\gamma/\Delta\mu$  (ref.<sup>31</sup>), where  $\Omega$  is the molecular volume,  $\Delta\mu = k_B T \ln(c_H/c_e)$  is the chemical potential difference between the solution and the crystal,  $k_B$  is the Boltzmann constant,  $T$  is the temperature,  $c_H$  is the haematin concentration and  $c_e$  is the solubility. In turn, lower  $\gamma$  and  $R_c$  stimulate

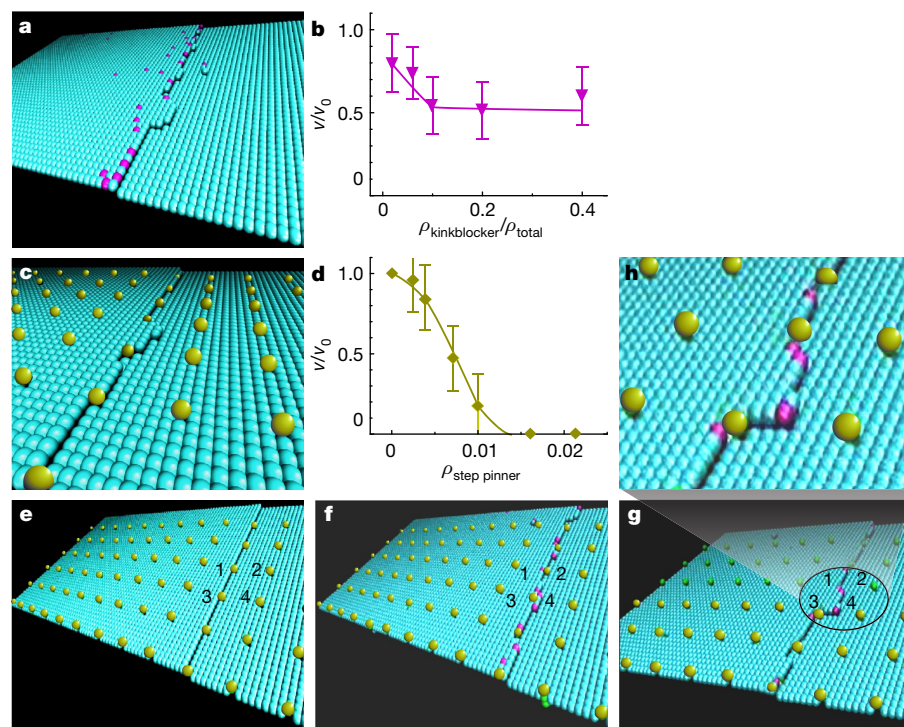




**Fig. 3 | Characterization of the effects of the kink blockers MQ and AQ on layer nucleation.** **a**, Time-resolved in situ AFM images showing growing (I and II) and dissolving (III) islands on a (100) face at  $c_H = 0.28$  mM and supersaturation  $\sigma = \ln(c_H/c_e) \approx 0.56$ . **b**, **c**, Dependence of the radius of the critical two-dimensional nucleus  $R_c$  on the crystallization driving force  $\Delta\mu = k_B T \ln(c_H/c_e)$  in pure haematin solution and in the presence of MQ (**b**) and AQ (**c**). Error bars represent the standard deviation from the average of 25 to 30 measurements. Solid lines are plots of the Gibbs–Thomson relation  $R_c = \Delta\gamma/\Delta\mu$  with step line tension  $\gamma = 25$  mJ m $^{-2}$  for pure haematin and 20 and 22 mJ m $^{-2}$  for MQ and AQ, respectively. Data for pure haematin are from Olafson et al.<sup>12</sup>

faster layer nucleation as  $J_{2D} = J_0 \exp[-\pi\gamma R_c h/(k_B T)]$  ( $h = 1.2$  nm is the step height)<sup>12</sup> and expedite step propagation in the gaps between the adsorbed step pinners (Fig. 1b). We developed (Supplementary Sections 3, 4, Extended Data Figs. 5, 6) an analytical model of the combined action of step pinners and kink blockers on step propagation and analysed the consequences of the presence of two types of inhibitors on the nucleation of a new crystal layer (Supplementary Section 5). This examination advocates that the classical synergistic effects dominate at low concentrations of either inhibitor, whereas the proposed mechanism of antagonism mobilizes at high concentrations; stronger antagonism between step pinners and kink blockers is projected for their joint action on  $J_{2D}$  rather than on  $\nu$  (Supplementary Section 5). Both predictions are borne by the  $J_{2D}$  and  $\nu$  correlations (Fig. 2b–e).

Data on layer nucleation in the presence of MQ or AQ demonstrate that  $\gamma$  decreases in the presence any of these inhibitors and that the measured  $\Delta\gamma$  correlates with the inhibition of step motion due to association of these inhibitors with the kinks. From AFM images, we directly measured  $R_c$  in the presence of 2.5  $\mu$ M MQ or AQ. This parameter represents the critical size of a two-dimensional nucleus of a crystal layer below which nuclei tend to dissolve, whereas nuclei larger than  $R_c$  have a greater probability to grow (Fig. 3a). We monitored the evolution of 25 to 30 layer nuclei at each value of  $\Delta\mu$  and inhibitor concentration, where  $\Delta\mu$  was varied by the selection of the haematin concentration  $c_H$  (Fig. 3b). The relation between  $R_c$  and  $\Delta\mu$  (Fig. 3c, d) is reciprocal, consistent with the Gibbs–Thomson relation, and reveals that the presence of MQ and AQ lowers  $\gamma$  from a nominal value of  $25 \pm 2$  mJ m $^{-2}$  to  $20 \pm 2$  and  $22 \pm 1$  mJ m $^{-2}$ , respectively. In Methods, we discuss statistical tests that certify the distinction of the three  $\gamma$  values and relate decreasing  $\gamma$  to association of MQ and AQ with the kinks. We assume the two kink blockers adsorb to the steps following a Langmuir-type law. In Supplementary Sections 1, 2, we evaluate  $-\Delta\gamma$  using the Gibbs equation of adsorption,  $F = -d\gamma/d\mu_B$ , where  $F$  is the amount of inhibitor absorbed at kinks and  $\mu_B = \mu_{B0} + k_B T \ln c_B$  is the chemical potential of the kink blocker, MQ or AQ, at concentration  $c_B$  (ref.<sup>31</sup>). From these relations and Extended Data Fig. 4, Extended Data Tables 3, 4, we obtain  $\Delta\gamma \approx -3$  mJ m $^{-2}$  for both MQ and AQ, in good agreement with the values for these two inhibitors assessed from the  $R_c(\Delta\mu)$  correlations



**Fig. 4 | Solid-on-solid kinetic Monte Carlo modelling of the action of kink blockers and step pinners on step propagation.** **a**, Kink blockers (magenta spheres) associate with kinks and incorporate in the crystal. **b**, Dependence of the step velocity  $\nu$  relative to that in pure solution  $\nu_0$  on the concentration of kink blockers  $\rho_{\text{kink blocker}}$  relative to  $\rho_{\text{total}}$ , the summed concentration of solute and kink blockers. **c**, Step pinners (gold spheres) adsorb on the terraces between steps and enforce curved steps. **d**, Dependence of the step velocity  $\nu$  relative to that in pure haematin solution  $\nu_0$  on the surface density of step pinners,  $\rho_{\text{step pinner}}$ . Error bars in **b** and **d** represent the standard error of the simulations, evaluated as discussed in Methods. **e**, Step pinners adsorbed on the surface arrest step advancement. Four numbered step pinners mark the step location. **f**, **g**, Addition of kink blockers stimulates the growth of a step stalled by step pinners; **g** presents a later moment of the same simulation as in **f**. **h**, Magnified view of a step squeezed between stoppers 1 and 3, showing kink blockers associated with kinks in the growing step segment.

(Fig. 3b, c). These  $\Delta\gamma$  invoke an equivalent contraction of  $R_c$  (ref. 13). As a 20% decrease in  $R_c$  is equivalent to a 1.44-fold ( $1.2^2$ ) lowering of the surface coverage of adsorbed step pinners, and given that  $J_{2D}$  and  $\nu$  are highly sensitive functions of both  $c_{CQ}$  and  $c_{QN}$ , the decrease in  $\gamma$  elicits a disproportionately strong response of  $\nu$  and  $J_{2D}$ .

In situ AFM measurements were complemented with kinetic Monte Carlo simulations to test the generality of the proposed model of antagonistic cooperativity between two classes of crystallization inhibitors. We developed a solid-on-solid model for step growth<sup>32</sup>, in which molecules associate and dissociate from steps. For simplicity, we ignored surface diffusion on the terraces. The rate of solute association depends on the supersaturation, whereas the probability of detachment is dictated by the bonds a molecule forms with its neighbours (Supplementary Video 1). We assume that kink blocker adsorption and detachment are analogous to solute molecules, and that the relevant dynamics are governed by their concentration and the number and strength of the bonds at an adsorption site (we assume that two of the lateral bonds are stronger and that the remaining two are weaker than for the solute molecules). These assumptions lead to preferential binding to the kinks at steps (Fig. 4a) and constrained  $\nu$  (Fig. 4b, Supplementary Video 2). We assume that step pinners bind strongly to the crystal surface, but exhibit no interactions with crystal molecules parallel to that plane. The surface is decorated with a square array of step pinners and they remain static throughout the simulation (Fig. 4c, Supplementary Video 3); previous results have demonstrated that the step-pinner surface distribution has no effect on the step velocity<sup>32</sup>. Remarkably, the calculated correlations between  $\nu$  and inhibitor concentrations are akin to those observed experimentally for the kink blockers MQ and AQ, for which  $\nu$  levels off at around 50% inhibition (Fig. 4b), as well as the step pinners CQ and QN, which induce complete growth arrest at moderate inhibitor concentrations (Fig. 4d)<sup>13</sup>.

Combining step pinners at a concentration above the threshold for complete growth arrest (Fig. 4e) with kink blockers allows steps to advance through pinned sites, thereby re-establishing layer growth (Fig. 4f, g, Supplementary Videos 4, 5). The simulations reveal that at the microscopic level, the antagonistic cooperativity is due to the stabilization of step edge fluctuations by associating kink blockers. Steps overcome the pinner palisade by fluctuations that penetrate the gaps between the pinners (Fig. 4c, Supplementary Video 5)<sup>32</sup>. Closely spaced pinners suppress the extent and lifetime of the fluctuations and restrain step growth. The blockers bind to the kink-rich fingers embodying the fluctuations (Fig. 4h) and increase the fluctuation lifetime. At the macroscopic level, the stabilized fluctuations manifest as a decrease in  $\gamma$ . Indeed, an attenuated  $\gamma$  enforces shorter  $R_c$ , which, in turn, allows step progress between the pinners (Fig. 4f–h).

In summary, we present a mechanism of antagonistic cooperativity between crystallization inhibitors by which kink blockers attenuate the step line tension and facilitate step propagation through the palisade of step pinners. This mechanism may provide guidance in the search for suitable inhibitor combinations to control crystallization of pathological, biomimetic and synthetic materials. In a broader context, our results highlight modifier interactions mediated by the dynamics and structures on the crystal interface as a prime element of the regulation of the shapes and patterns of crystalline structures in nature and industry.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information,

acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-019-1918-4>.

- Marin, F. & Luquet, G. Molluscan biomineralization: the proteinaceous shell constituents of *Pinna nobilis* L. *Mater. Sci. Eng. C* **25**, 105–111 (2005).
- Porter, S. M. Seawater chemistry and early carbonate biomineralization. *Science* **316**, 1302 (2007).
- Myerson, A. S. *Handbook of Industrial Crystallization* 2nd edn (Butterworth-Heinemann, 2001).
- Sangwal, K. *Additives and Crystallization Processes: From Fundamentals to Applications* (John Wiley & Sons, 2007).
- De Yoreo, J. J. Physical mechanisms of crystal growth modification by biomolecules. *ALP Conf. Proc.* **1270**, 45–58 (2010).
- Orme, C. A. et al. Formation of chiral morphologies through selective binding of amino acids to calcite surface steps. *Nature* **411**, 775–779 (2001).
- De Yoreo, J. J. et al. Crystallization by particle attachment in synthetic, biogenic, and geologic environments. *Science* **349**, aaa6760 (2015).
- Elhadi, S., De Yoreo, J. J., Hoyer, J. R. & Dove, P. M. Role of molecular charge and hydrophilicity in regulating the kinetics of crystal growth. *Proc. Natl Acad. Sci. USA* **103**, 19237–19242 (2006).
- Ferrer, M. D. et al. A novel pharmacodynamic assay to evaluate the effects of crystallization inhibitors on calcium phosphate crystallization in human plasma. *Sci. Rep.* **7**, 6858 (2017).
- Pagola, S., Stephens, P. W., Bohle, D. S., Kosar, A. D. & Madsen, S. K. The structure of malaria pigment  $\beta$ -haematin. *Nature* **404**, 307–310 (2000).
- Sullivan, D. J., Matile, H., Ridley, R. G. & Goldberg, D. E. A common mechanism for blockade of heme polymerization by antimalarial quinolines. *J. Biol. Chem.* **273**, 31103–31107 (1998).
- Olafson, K. N., Ketchum, M. A., Rimer, J. D. & Vekilov, P. G. Mechanisms of hematin crystallization and inhibition by the antimalarial drug chloroquine. *Proc. Natl Acad. Sci. USA* **112**, 4946–4951 (2015).
- Olafson, K. N., Nguyen, T. Q., Rimer, J. D. & Vekilov, P. G. Antimalarials inhibit hematin crystallization by unique drug–surface site interactions. *Proc. Natl Acad. Sci. USA* **114**, 7531–7536 (2017).
- Van Horn, H. M. The crystallization of white dwarf stars. *Nat. Astron.* **3**, 129–130 (2019).
- Reznikov, N., Steele, J. A. M., Fratzl, P. & Stevens, M. M. A materials science vision of extracellular matrix mineralization. *Nat. Rev. Mater.* **1**, 16041 (2016).
- De Yoreo, J. J. & Vekilov, P. G. Principles of crystal nucleation and growth. *Rev. Mineral. Geochem.* **54**, 57–93 (2003).
- Olafson, K. N., Li, R., Alamani, B. G. & Rimer, J. D. Engineering crystal modifiers: bridging classical and nonclassical crystallization. *Chem. Mater.* **28**, 8453–8465 (2016).
- Rae Cho, K. et al. Direct observation of mineral–organic composite formation reveals occlusion mechanism. *Nat. Commun.* **7**, 10187 (2016).
- Weiner, S. & Addadi, L. Design strategies in mineralized biological materials. *J. Mater. Chem.* **7**, 689–702 (1997).
- Meldrum, F. C. & Cölfen, H. Controlling mineral morphologies and structures in biological and synthetic systems. *Chem. Rev.* **108**, 4332–4432 (2008).
- Farmanesh, S. et al. Specificity of growth inhibitors and their cooperative effects in calcium oxalate monohydrate crystallization. *J. Am. Chem. Soc.* **136**, 367–376 (2014).
- Ridley, R. G. Medical need, scientific opportunity and the drive for antimalarial drugs. *Nature* **415**, 686–693 (2002).
- Sullivan, D. J., Gluzman, I. Y., Russell, D. G. & Goldberg, D. E. On the molecular mechanism of chloroquine's antimalarial action. *Proc. Natl Acad. Sci. USA* **93**, 11865–11870 (1996).
- Gorka, A. P., de Dios, A. & Roepe, P. D. Quinoline drug–heme interactions and implications for antimalarial cytostatic versus cytotoxic activities. *J. Med. Chem.* **56**, 5231–5246 (2013).
- Cabrera, N. & Vermilyea, D. A. in *Growth and Perfection of Crystals* Vol. **393–408** (eds Doremus, R. H. et al.) 393–408 (Wiley, 1958).
- Eastman, R. T. & Fidock, D. A. Artemisinin-based combination therapies: a vital tool in efforts to eliminate malaria. *Nat. Rev. Microbiol.* **7**, 864–874 (2009).
- Mott, B. T. et al. High-throughput matrix screening identifies synergistic and antagonistic antimalarial drug combinations. *Sci. Rep.* **5**, 13891 (2015).
- Gorka, A. P., Jacobs, L. M. & Roepe, P. D. Cytostatic versus cytotoxic profiling of quinoline drug combinations via modified fixed-ratio isobologram analysis. *Malar. J.* **12**, 332 (2013).
- Chou, T.-C. Theoretical basis, experimental design, and computerized simulation of synergism and antagonism in drug combination studies. *Pharmacol. Rev.* **58**, 621–681 (2006).
- Egan, T. J. Interactions of quinoline antimalarials with hematin in solution. *J. Inorg. Biochem.* **100**, 916–926 (2006).
- Gibbs, J. W. On the equilibrium of heterogeneous substances (first part). *Trans. Connect. Acad. Sci.* **3**, 108–248 (1876).
- Lutsko, J. F. et al. Crystal growth cessation revisited: the physical basis of step pinning. *Cryst. Growth Des.* **14**, 6129–6134 (2014).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2020

### Solution preparation

The following compounds were purchased from Sigma Aldrich: haematin porcine ( $\geq 98\%$ ), citric acid (anhydrous,  $\geq 99.5\%$ ), sodium hydroxide (anhydrous,  $\geq 98\%$ ), *n*-octanol (anhydrous,  $\geq 99\%$ ), porcine haematin, chloroquine diphosphate ( $\geq 98\%$ ), quinine (anhydrous,  $\geq 98.0\%$ ), amodiaquine dihydrochloride dihydrate and mefloquine hydrochloride (anhydrous,  $\geq 98.0\%$ ). All reagents were used as received. Deionized water was produced by a Millipore reverse osmosis–ion exchange system (Rios-8 Proguard 2–MilliQ Q-guard).

Citric buffer at pH 4.80 was prepared by dissolving 50 mM of citric acid in deionized water and titrating the solution, under continuous stirring, with 0.10 M NaOH to the desired pH. The buffer pH was verified before each experiment and fresh buffers were prepared every month. We placed 5 ml of citric buffer at pH 4.80 in direct contact with *n*-octanol at 23 °C and allowed 30 min for equilibration. The upper portion of the two-phase system was decanted and denoted as citric buffer-saturated octanol (CBSO).

For this study, we used four antimalarial drugs: QN, CQ, AQ and MQ. Solid QN and MQ were added to CBSO and the solutions reached the desired concentration after 2–4 d. AQ and CQ were added in excess to CBSO and stored in the dark for 30–45 d, allowing the concentrations to approach the respective solubilities<sup>33</sup>. All drug solutions were filtered through 0.2  $\mu\text{m}$  nylon membrane filters and the concentrations were determined by ultraviolet–visible spectrometry using a Beckman DU 800 spectrophotometer and extinction coefficients and wavelengths listed in Ketchum et al.<sup>33</sup>.

Haematin solutions were prepared by dissolving haematin powder in 8 ml of freshly made CBSO and heating it to 70 °C for 7–9 h. The solution was filtered through a 0.2  $\mu\text{m}$  nylon membrane filter and the concentration was determined using an extinction coefficient of  $3.1 \pm 0.1 \text{ cm}^{-1} \text{ mM}^{-1}$  at a wavelength of 594 nm (refs. <sup>12,34</sup>).

### Characterization of the combined inhibitor effects on bulk haematin crystallization

We adopted the procedure reported by Olafson et al.<sup>12,35</sup> to produce haematin crystals from supersaturated haematin solution in CBSO. We tested crystal growth in the presence of four drug combinations, CQ/MQ, CQ/AQ, QN/MQ and QN/AQ, with constant ratios between the two constituents of 1:4, 1:2, 1:2 and 1:2, respectively. Drug combinations were added to the haematin stock solution to achieve final total inhibitor concentrations ranging from 0 to 15  $\mu\text{M}$  while maintaining a constant haematin concentration ( $c_{\text{H}} = 0.28 \text{ mM}$ ). The vials were then shaken until the solution was well mixed. A 15- $\mu\text{m}$ -diameter glass slide was scratched in the centre and placed at the bottom of the vial in contact with the supersaturated solution. Vials were capped and placed in an incubator at 23 °C with minimal exposure to light.  $\beta$ -Haematin crystals were observed in 1–2 d and reached their maximum length after around 2 weeks. The glass slide with attached haematin crystals was collected, washed with deionized water and ethanol, dried with nitrogen gas and then coated with 10–20 nm gold for scanning electron microscopy. The length and width of about 30 crystals at each composition were measured to assess the effectiveness of inhibitor combinations.

### In situ monitoring of the haematin crystal evolution

We used a multimode atomic force microscope (Nanoscope IV) from Digital Instruments for all AFM experiments. AFM images were collected in tapping mode using Olympus TR800PSA probes (silicon nitride, Cr/Au coated 5/30, 0.15 N m<sup>-1</sup> spring constant) with a tapping frequency of 32 kHz. Image sizes ranged from 300 nm to 20  $\mu\text{m}$ . Scan rates were between 1 and 2.52 s<sup>-1</sup>. Height, amplitude and phase imaging modes were employed. The captured images contained 256 scan lines at angles depending on the orientation of the monitored crystal<sup>12,36</sup>.

The temperature in the fluid cell reached a steady value of  $27.8 \pm 0.1$  °C within 15 min of imaging<sup>36</sup>. This value was higher than room temperature (around 22 °C) owing to heating by the AFM scanner and laser.

$\beta$ -Haematin crystals were grown on glass disks as described above. The density of attached haematin crystals was monitored under an optical microscope. We ensured similar crystal density for all samples to minimize potential depletion of inhibitors due to high crystal number. The glass slides were mounted on AFM sample disks (Ted Pella) and the samples were placed on the AFM scanner. Haematin solution in CBSO with a concentration of 0.28 mM was prepared less than 2 h in advance. This solution was loaded into the AFM liquid cell using 1 ml disposable polypropylene syringes (Henck Sass Wolf), tolerant of organic solvents. After loading, the system was left standing for 10–20 min to thermally equilibrate. The crystal edges were identified to determine the orientation and the crystallographic directions on the upward-facing (100) crystal surface. The crystals were kept in contact with the solution for 0.5–1.5 h to allow their surface features to adapt to growth conditions. We set the scan direction parallel to the [001] crystallographic direction and AFM images were collected for 3–5 h. The solution in the AFM fluid cell was refreshed every 30 min to maintain constant concentration. For studies of modifiers, growth solutions were replaced with ones containing a selected antimalarial inhibitor(s). With each modifier concentration, AFM images were collected for 2 to 4 h, during which the solution was replenished several times. Solution without modifier was pumped into the AFM cell and the observed crystal was allowed to grow uninhibited for about 30 min before another modifier concentration was introduced.

The evolution of the haematin crystal surface was characterized by the velocity of growing steps  $v$  and the rate of two-dimensional nucleation of new crystal layers  $J_{2D}$ . To determine  $v$ , we monitored the displacements of 8–13 individual steps with a measured step height  $h = 1.17 \pm 0.07 \text{ nm}$ . Between 25 and 35 measurements were taken for each individual step and the average growth rates were reported. To determine  $J_{2D}$ , the appearance of new islands on the surface between successive images was monitored and the number of islands that grew was counted. This number was scaled with the imaged area and the time interval between images to yield  $J_{2D}$ . From 15 to 25 measurements were averaged for each  $J_{2D}$  determination.

The goal of the AFM investigations was to establish the molecular mechanisms of synergy or antagonism between step pinners and kink blockers in inhibiting the growth of  $\beta$ -haematin crystals. Using AFM imaging at the mesoscopic scale, we demonstrate that step pinners and kink blockers cooperate in suppressing both the nucleation of new layers and the propagation of steps on haematin crystal surfaces. The nucleation of new layers at random locations on the crystal surface requires observations at the mesoscopic length scale, within the range of capabilities of standard AFM techniques. Images with molecular resolution of growing steps would have provided additional insights. As shown in our previous work on haematin crystallization, imaging with resolution comparable to the size of the haematin molecule, around 1 nm, is possible during in situ AFM monitoring of flat crystal planes<sup>13</sup>. The presence of steps, however, disrupts the contact between the scanning tip and the crystal surface and lowers the image resolution. Strict numerical correspondence between discrete molecular-level events and the mesoscopic and macroscopic variables that characterize crystal growth and inhibition has been established in our earlier work<sup>37–40</sup>. This correspondence supports the molecular mechanisms based on observations at mesoscopic length scales.

### Determination of the surface free energy of the step edge $\gamma$ in the presence of MQ and AQ

We evaluate the value of  $\gamma$  from the correlation between the radius of the two-dimensional nucleus of new layers  $R_c$  and the supersaturation, similar to previous determinations in solutions without inhibitors carried out by Olafson et al.<sup>12</sup>. The critical radius  $R_c$  for layer nucleation

is defined as the threshold size above which an island has a higher probability to grow. Islands of size  $R < R_{\text{crit}}$  are more likely to dissolve. We monitored the size evolution of all newly generated islands from time-resolved sequences of in situ AFM images and classified the islands as growing or dissolving. The largest sizes reached by dissolving islands and the threshold, above which all islands grew, were averaged to yield  $R_c$ . We determined from 25 to 30 independent  $R_c$  measurements at each combination of haematin and MQ or AQ concentration. Six concentrations of haematin  $c_H$  were tested in the presence of 2.5  $\mu\text{M}$  MQ and seven in the presence of 2.5  $\mu\text{M}$  AQ. The  $R_c$  values obtained at each concentration of the two inhibitors were averaged and plotted as a function of the supersaturation  $\Delta\mu = k_B T \ln(c_H/c_e)$ , and were compared with the values of  $R_c$  in the absence of inhibitors (Fig. 3b, c).

The Gibbs–Thomson relation  $R_c = \Omega\gamma/\Delta\mu$ , where  $\Omega = 0.708 \text{ nm}^3$  is the molecular volume in the crystal, prescribes the values of  $\gamma$  corresponding to each of the  $R_c(\Delta\mu)$  correlations:  $25 \pm 2 \text{ mJ m}^{-2}$  in solution without inhibitors,  $20 \pm 2 \text{ mJ m}^{-2}$  in the presence of MQ and  $22 \pm 1 \text{ mJ m}^{-2}$  in the presence of AQ. The standard deviations of the three  $\gamma$  values arise from the regression analyses of the linear correlations  $R_c(\Delta\mu)^{-1}$  and reveal that the confidence intervals of  $\gamma$  at the three tested solution compositions partially overlap.

We analysed the similarity between the three individual values of  $\gamma$  by one-way analysis of variance, a statistical procedure that compares the variance between two groups to the variance within each group of data. We computed individual  $\gamma$  values from each  $R_c$  measurement and examined the similarity between three pairs of  $\gamma$  datasets: no inhibitor/AQ, no inhibitor/MQ and MQ/AQ. The analysis of variance test parameters are listed in Extended Data Table 2. The three  $F$  values, corresponding to the ratio of the variances within each pair of datasets, are significantly greater than the critical values for groups consisting of 195, 177 and 297 independent measurements. The  $P$  values were of the order of  $10^{-3}$ ,  $10^{-6}$  and  $10^{-7}$ , respectively, smaller than the significance level of 0.05. These  $F$  and  $P$  values consonantly certify that the hypothesis of equality of the three  $\gamma$  values is rejected.

### Inhibitor–inhibitor complexation

The aim of these tests was to determine whether binary complexes between paired inhibitors form and reduce the inhibitor concentrations. Spectroscopic characterization of solutions of the tested inhibitors revealed that the sum of the ultraviolet-visible absorbances of individual inhibitors is approximately identical to the absorbance of their combination. (Extended Data Fig. 3a–d). Moreover, no shift in absorbance peaks was observed after mixing. These results suggest that it is unlikely that complexes form between two inhibitors.

### Inhibitor–haematin–inhibitor complexation

Complexes formed between haematin and antimalarial inhibitors have been discussed by Egan and co-workers<sup>30,41</sup> and the complexation constants between haematin and antimalarial inhibitors in CBSO have been reported by Olafson et al.<sup>13</sup> Using established protocols, we tested for the complexation between haematin and four inhibitor pairs: QN/AQ, QN/MQ, CQ/AQ and CQ/MQ. The two tested inhibitors were dissolved at equal concentrations in CBSO and 2 ml of this stock was mixed to a final concentration determined by the lower inhibitor solubility. Fresh haematin stock was diluted with CBSO to a concentration of 0.38 mM and then titrated with a solution of the inhibitor pair. At each titration step, the solution was gently stirred for 8–10 min to complete complexation and a 350  $\mu\text{l}$  aliquot was drawn for ultraviolet-visible spectrometry. The ultraviolet-visible absorptions at 594 nm were measured for 40 titration steps and rescaled to account for the dilution. The rescaled absorbance  $A_{\text{corr}}$  was compared with a theoretical curve calculated from the complexation constants of the two tested inhibitors.

The absorbance at around 594 nm displayed a clear shift to higher wavelengths after the addition of the inhibitor mixture, which indicates the formation of complexes. We calculated the theoretical  $A_{\text{corr}}/A_0$

values (where  $A_0$  is the absorbance of a pure haematin solution) for four different models for each combination and chose the best fit from the minimal mean squared deviation between experimental and theoretical  $A_{\text{corr}}/A_0$  values. Non-zero deviations suggest the formation of complexes. The ultraviolet-visible spectra of solutions containing two inhibitors and haematin indicate that in all four combinations, even if new complexes exist, their concentration would be limited to a level that does not appreciably attenuate the concentration of antimalarial inhibitors in solution (Extended Data Fig. 3e–i). Therefore, the sequestration of inhibitors due to the formation of ternary inhibitor–haematin–inhibitor complexes is unlikely to be significant.

### Kinetic Monte Carlo model of cooperativity between step pinners and kink blockers

We employ a standard solid-on-solid kinetic Monte Carlo (kMC) model of crystal growth. We use a surface of a Kossel crystal consisting of  $N_x = 50$  by  $N_y = 100$  sites occupied by  $N = 5,000$  surface molecules. In the kMC algorithm, a surface site is chosen at random and one of the possible kMC actions is performed on the basis of the probability of the various actions;  $N$  repetitions of this act comprise one kMC time step. In the absence of inhibitors, three actions are possible at a surface site: a molecule attaches to the site, the molecule occupying the site detaches or nothing happens (that is, the molecule remains fixed). The probability for attachment is  $dt \times v e^{\mu/(k_B T)}$ , where  $dt$  is the kMC time step,  $v$  is the inverse kMC timescale and  $\mu$  is the chemical potential. The probability for a molecule to detach from site  $i$  is  $dt \times v e^{E_i/(k_B T)}$ , where  $E_i$  is the energy of the surface molecule at site  $i$ . The energy  $E_i$  is evaluated as the sum of the bond energies of the molecule with its six nearest neighbours. In a pure crystal, the bond energy is taken to be the same in all directions and is denoted  $\epsilon$ . By expressing temperatures in the dimensionless form  $k_B T/\epsilon$ , the physical value of  $\epsilon$  is not needed. Given that a molecule in the bulk crystal has six bonds with the energy shared between it and its neighbours, the binding energy in the bulk is  $3\epsilon$  per molecule and so the equilibrium chemical potential is  $\mu_{\text{equil}} = 3\epsilon$ .

Inhibitors are handled in two distinct ways. Static inhibitors function as step pinners. They are deposited on the surface at the beginning of a simulation and do not participate in the kMC actions. When a crystal molecule is next to a static pinner, the bond energy between the two is taken to be zero. Thus, the only parameter needed to characterize the pinners is their surface density. As they do not contribute to the binding of molecules to the crystal, the pinners disrupt and impede the growth of surface layers. For conceptual simplicity, we arranged the pinners in a square grid (Fig. 4c). If the pinners are too close together (that is, if their surface density is too high), the step velocity is zero and crystal growth is arrested. The physics of step blocking by such inhibitors, the criterion for step pinning, and a demonstration that inhibition is independent of the physical arrangement of the step blockers has been extensively discussed in Lutsko et al.<sup>32</sup>

A new feature of the present simulation work is the model of kink blockers. Similar to the solute molecules, the kink blockers are dynamic. In the presence of kink blockers, the pool of possible events at a crystal site is expanded to include their attachment and detachment. To block the kinks, the kink blockers must differ from the solute species and from the step pinners. We assume, for simplicity, that kink blockers do not bind to step pinners. We also assume that the kink blockers bind to the molecules in the crystal with a non-zero binding energy, otherwise, they would not exhibit a preference for kink sites. The kink blocker can only impede step growth if the bonding is weaker than the intermolecular bonds in the crystal  $\epsilon$ . In contrast, weakly bound inhibitors would have a low residence time at the kinks and have little or no effect on step growth<sup>32</sup>. To reconcile these two requirements, we assume that the kink blockers bind anisotropically. We assume that the only non-zero bonds formed by kink blockers are to in-plane crystal molecules. Furthermore, we assume that the in-plane bond strengths are not equal. Two out of the four in-plane bonding directions are randomly assigned



bond strength  $2\epsilon$  and the other two,  $0.5\epsilon$ . The bond energy of a solute molecule deposited on top of a kink blocker is  $\epsilon$ .

The total energy of a kink blocker surrounded by crystal molecules is  $6\epsilon$ , equal to the crystal molecules so that the incorporation of kink blockers does not change the energetics of crystal growth. However, the asymmetry of their binding to the crystal surface modifies the kinetics of step growth. A kink blocker attached to a kink site with orientation that promotes two bonds of total energy  $4\epsilon$  will be bound stronger than a solute molecule bound with energy  $3\epsilon$ . Such kink blockers are unlikely to detach. In contrast, the bonds that this kink blocker molecule can form with the incoming solute molecules are weak and solute molecules that deposit next to it are more likely to detach than if deposited in a free kink. These dynamics impede step growth. A kink blocker attached to a kink in an unfavourable orientation, or adsorbed at a non-kink surface site, would have a total energy of  $2.5\epsilon$  or less and will tend to detach.

Our kMC model is subject to several constraints. First, the only model parameters that one can easily vary are the bond strengths in the various directions. Second, a foreign molecule acts as a kink blocker if (1) it is attracted to kink sites, (2) it inhibits step growth and (3) it has a sufficient residency time to affect the step growth dynamics. These requirements inevitably lead to asymmetric lateral bonds with a total binding energy in a kink site equal to or greater than the energy of a crystal molecule in a kink site. Within these constraints, we do not expect our results to strongly depend on the numerical values chosen.

Errors were estimated by averaging the step velocity over windows of 1,000 surface updates, thus producing a set of independent estimates of the velocity during the simulations. The arithmetic average of these values gives the overall estimate of the step velocity and the root-mean-squared deviation from the average of the averages is used to estimate its standard deviation. The error bars reported in the figures are the standard errors of the step velocities calculated as their standard deviations divided by the square root of the number of samples.

## Data availability

The datasets generated during and/or analysed during the current study are available from the corresponding authors on reasonable request.

## Code availability

The custom computer code used in these simulations is available upon reasonable request to J.F.L. (jim@lutsko.com).

33. Ketchum, M. A., Lee, A. M., Vekilov, P. G. & Rimer, J. D. Biomimetic assay for hematin crystallization inhibitors: a new platform to screen antimalarial drugs. *Cryst. Growth Des.* **17**, 197–206 (2017).
34. Ketchum, M. A., Olafson, K. N., Petrova, E. V., Rimer, J. D. & Vekilov, P. G. Hematin crystallization from aqueous and organic solvents. *J. Chem. Phys.* **139**, 121911 (2013).
35. Olafson, K. N., Rimer, J. D. & Vekilov, P. G. Growth of large hematin crystals in biomimetic solutions. *Cryst. Growth Des.* **14**, 2123–2127 (2014).
36. Olafson, K. N., Ketchum, M. A., Rimer, J. D. & Vekilov, P. G. Molecular mechanisms of hematin crystallization from organic solvent. *Cryst. Growth Des.* **15**, 5535–5542 (2015).
37. Yau, S.-T., Thomas, B. R. & Vekilov, P. G. Molecular mechanisms of crystallization and defect formation. *Phys. Rev. Lett.* **85**, 353–356 (2000).
38. Yau, S.-T. & Vekilov, P. G. Quasi-planar nucleus structure in apoferritin crystallisation. *Nature* **406**, 494–497 (2000).
39. Yau, S.-T., Thomas, B. R., Galkin, O., Gliko, O. & Vekilov, P. G. Molecular mechanisms of microheterogeneity-induced defect formation in ferritin crystallization. *Proteins* **43**, 343–352 (2001).
40. Petsev, D. N., Chen, K., Gliko, O. & Vekilov, P. G. Diffusion-limited kinetics of the solution-solid phase transition of molecular substances. *Proc. Natl Acad. Sci. USA* **100**, 792–796 (2003).
41. Egan, T. J. & Ncoke, K. K. Effects of solvent composition and ionic strength on the interaction of quinalone antimalarials with ferriprotoporphyrin IX. *J. Inorg. Biochem.* **98**, 144–152 (2004).

**Acknowledgements** We thank K. Olafson for help with haematin crystallization and AFM analysis, D. Sullivan for discussions on haemozoin formation and drug–haematin interactions, and D. Maes for insights on experiment statistics. This work was supported by the National Science Foundation (award number DMR-1710354), the National Institutes of Health (award number 1R21AI126215-01), NASA (award numbers NNX14AD68G and NNX14AE79G), the European Space Agency (ESA) and the Belgian Federal Science Policy Office (BELSPO) in the framework of the PRODEX Programme (contract number ESA17 AO-2004-070) and The Welch Foundation (grant E-1794).

**Author contributions** J.D.R. conceived this work, P.G.V. and J.D.R. designed the experiments, W.M. performed all experiments, P.G.V. and W.M. analysed data, P.G.V. developed interpretive models, J.F.L. carried out the kMC simulations, and P.G.V., J.F.L. and J.D.R. wrote the paper. All authors discussed the results and commented on the manuscript.

**Competing interests** The authors declare no competing interests.

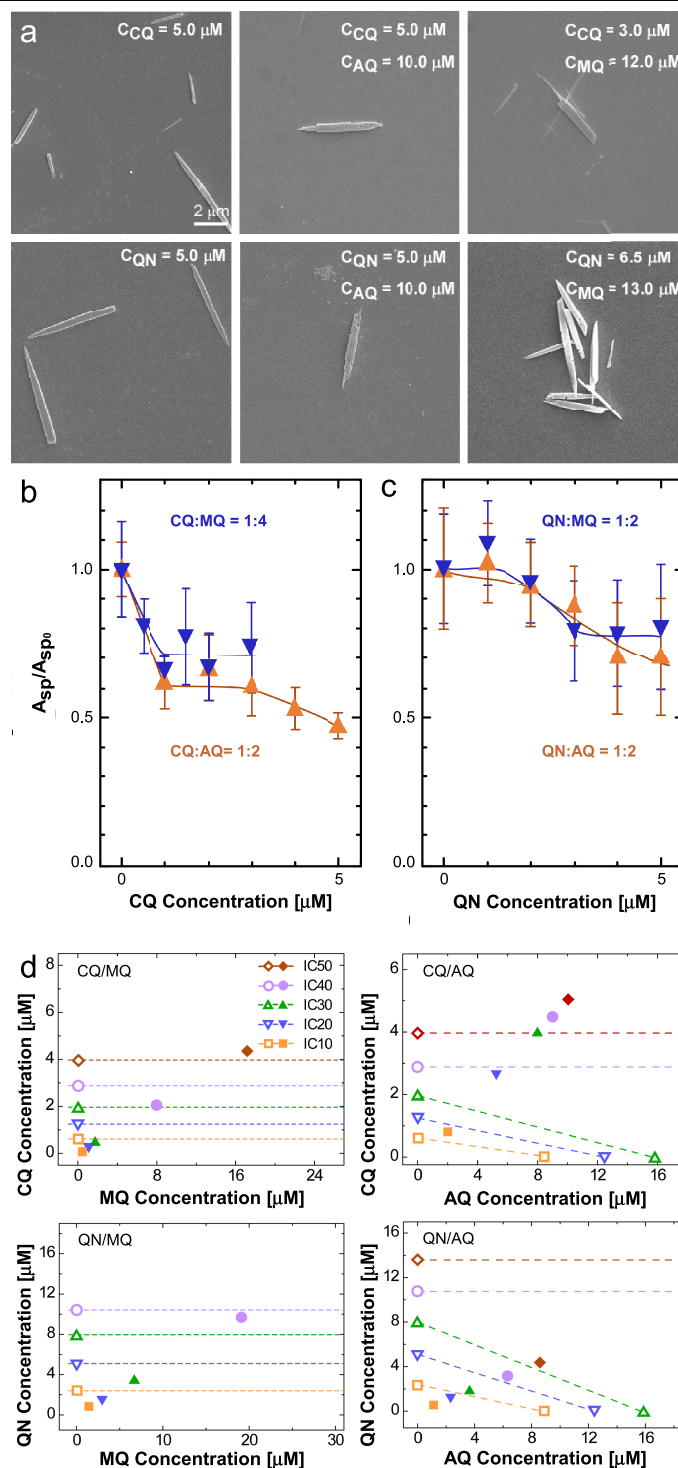
## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41586-019-1918-4>.

**Correspondence and requests for materials** should be addressed to J.D.R. or P.G.V.

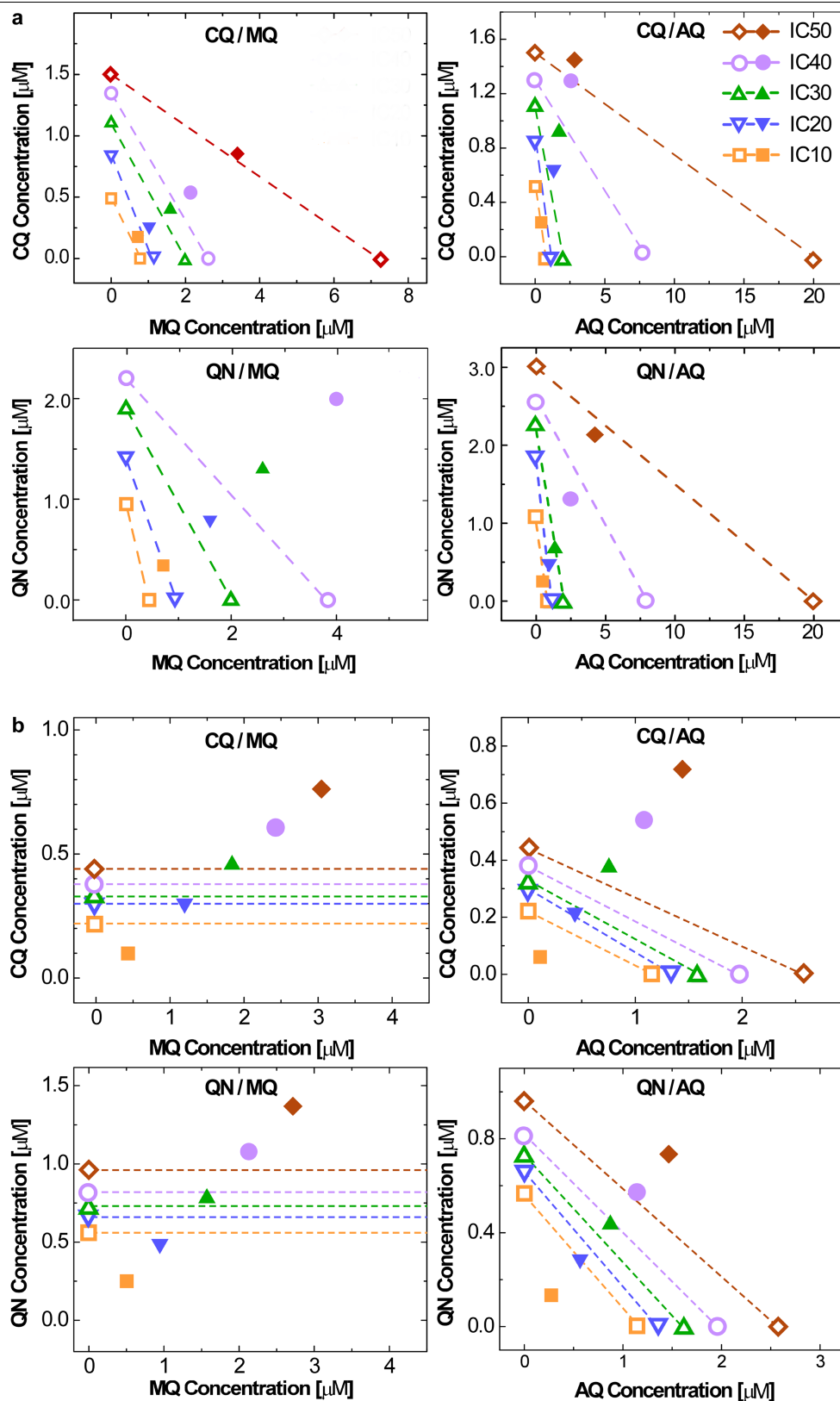
**Peer review information** Nature thanks Baron Peters and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>.



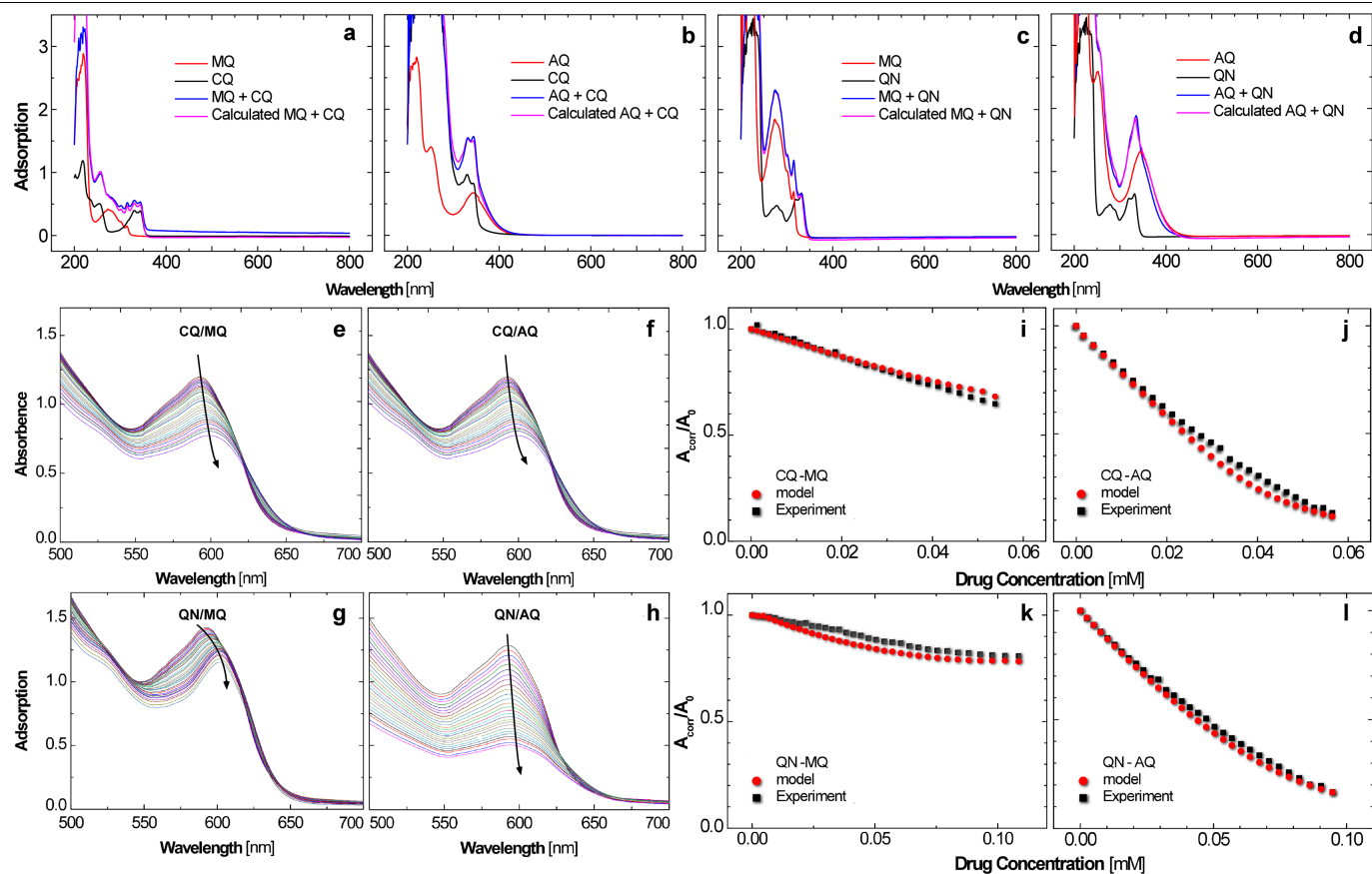
**Extended Data Fig. 1 | Effects of step pinners and kink blockers on bulk haematin crystallization.** **a**, Scanning electron microscopy micrographs of crystals grown in the presence of inhibitors at the concentrations listed in each panel for 16 d at 23 °C. **b**, **c**, Variations of the average length-to-width,  $l/w$ , aspect ratio  $A_{sp}$  of crystals grown in the presence of increasing concentrations of CQ/MQ and CQ/AQ (**b**) and QN/MQ and QN/AQ (**c**) at the displayed ratios relative to the  $A_{sp}$  reached after growth in pure CBSO solutions for 16 d at 23 °C. Lines are guides for the eye. Variations of the corresponding average crystal length  $l$  and width  $w$  are displayed in Fig. 1f–i. **d**, Isobolograms characterizing the cooperativity of the CQ/MQ, CQ/AQ, QN/MQ and QN/AQ inhibitor pairs in

suppressing the length of  $\beta$ -haematin crystals. Open symbols indicate the concentrations of individual inhibitors that elicit a certain percentage of inhibition, referred to as ICs. Dashed lines correspond to additive cooperativity between the paired inhibitors for a certain percentage of inhibition and are horizontal if the inhibitor in the abscissa is inactive when applied alone. Solid symbols represent the concentrations of the paired inhibitors that evoke the same inhibition. Rightward shifts of the solid symbols from the respective dashed lines indicate antagonistic cooperativity. The corresponding combination index values are listed in Extended Data Table 1.



**Extended Data Fig. 2 | Isobolograms characterizing the cooperativity of the CQ/MQ, CQ/AQ, QN/MQ and QN/AQ inhibitor pairs. a, b,** Cooperativity in suppressing the step velocity  $v$  (**a**) and the rate of two-dimensional nucleation rate  $J_{2D}$  of new layers (**b**). Open symbols indicate the concentrations of individual inhibitors that elicit a certain percentage of inhibition (ICs). Dashed lines correspond to additive cooperativity between the paired inhibitors for a

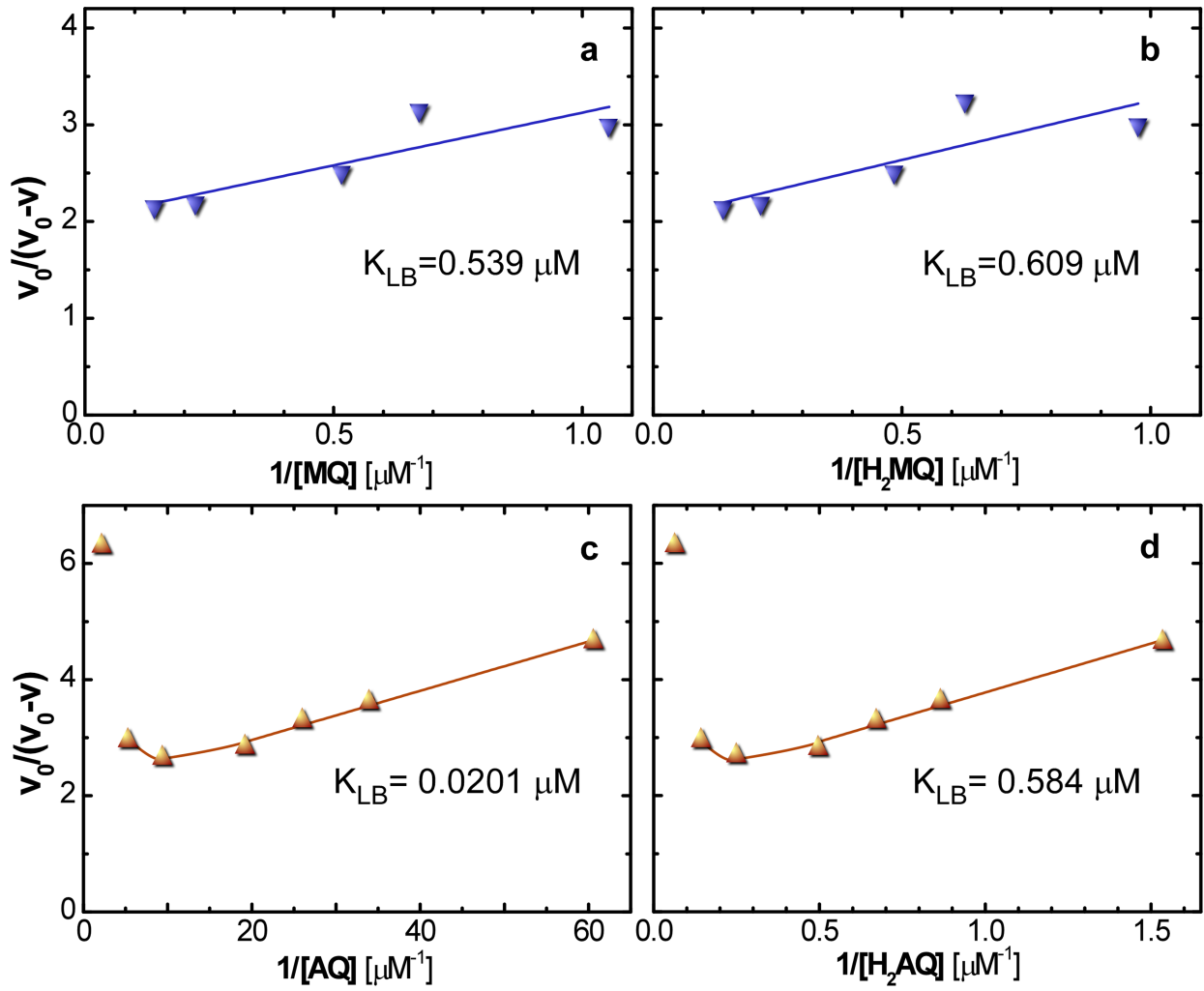
certain percentage of inhibition and are horizontal if the inhibitor in the abscissa is inactive when applied alone. Solid symbols represent the concentrations of the paired inhibitors that evoke the same inhibition. Rightward shifts of the solid symbols from the respective dashed lines indicate antagonistic cooperativity. The corresponding combination index values are listed in Extended Data Table 1.



**Extended Data Fig. 3 | Lack of complexation between kink blockers and step pinners in the solution. a–d,** Lack of CQ/MQ (a), CQ/AQ (b), QN/MQ (c) and QN/AQ (d) complexes. The ultraviolet-visible absorption spectra of the individual inhibitors and binary combinations indicated in the plots. The spectra of the binary solutions are nearly identical to the sum of the spectra of the individual inhibitors. **e–h,** Lack of ternary compounds that include haematin and the CQ/MQ (e, i), CQ/AQ (f, j), QN/MQ (g, k) and QN/AQ (h, l) pairs of inhibitors. **e–h,** The ultraviolet-visible spectra of haematin at concentration

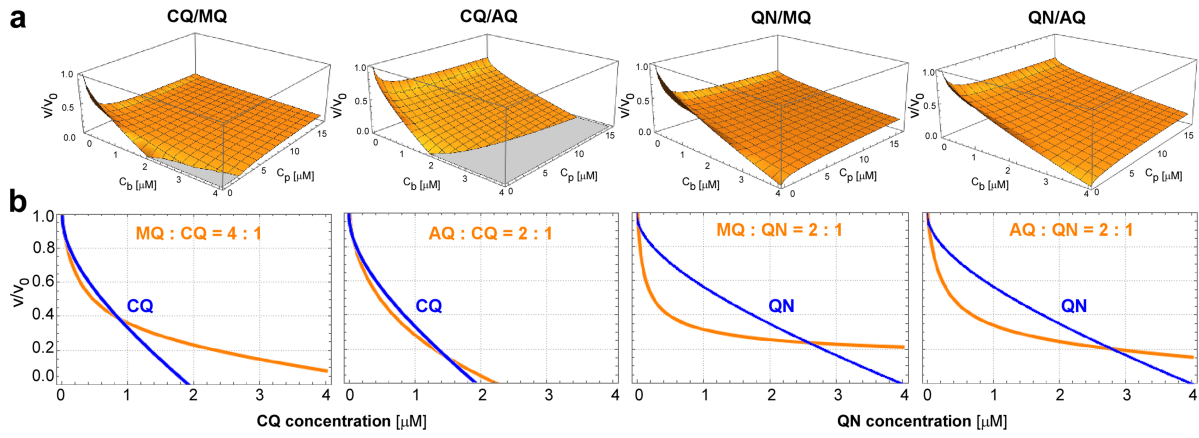
$c_H = 0.38$  mM in the presence of various combinations of QN, CQ, AQ and MQ (as indicated in the plots) at 1:1 molar ratios, where the inhibitor concentrations increase from top to bottom, as indicated by arrows. **i–l,** The relative decrease of the absorbance of a solution with initial  $c_H = 0.38$  mM at 594 nm as a function of the concentration of the respective inhibitor pair (1:1 ratio) compared with a model assuming the presence of complexes of haematin with each of the individual inhibitors in the mixture, evaluated using the haematin-inhibitor binding constants from Olafson et al.<sup>13</sup>.





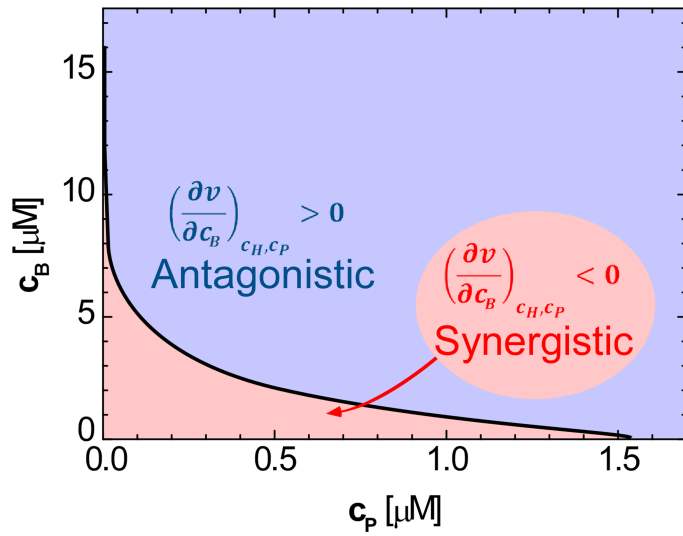
**Extended Data Fig. 4 | The correlation between the step velocity  $v$  and the inhibitor concentration.** a–d, Data are presented in linearized coordinates  $v_0/(v_0-v)^{-1}$  and  $c_B^{-1}$  ( $c_B$ , kink blocker concentration) according to Supplementary equation (7), for  $c_B = [B]$  (a, c) and  $c_B = [H_2B]$  (b, d), respectively, for MQ (a, b) and AQ (c, d). Original data on the dependence of the step velocity on the concentration of the kink blockers MQ and AQ are from Olafson et al.<sup>13</sup>.

The values of the Langmuir constant  $K_{LB}$  determined from the slope of the straight lines are shown. The two leftmost data points for AQ, measured at  $C_{AQ} > 7 \mu M$ , correspond to an unphysical increase in  $v$  at increasing concentration of AQ and were not considered in the regression analysis to determine  $K_{LB}$ .



**Extended Data Fig. 5 | The step velocity  $v$  in the presence step pinners and kink blockers, relative to that in pure solutions  $v_0$ .** Data calculated using Supplementary equation (22). The values of  $\xi$  and  $K_{LB}$  are listed in Extended Data Table 4.  $\gamma_0 = 25 \text{ mJ m}^{-1}$  is evaluated from the  $R_c$  determinations in Fig. 3.  $K_{LP} = 0.0027 \text{ } \mu\text{M}^{-1}$  for CQ and  $0.0013 \text{ } \mu\text{M}^{-1}$  QN is evaluated from the  $v(c_p)$  correlations for CQ and QN determined by Olafson et al.<sup>13</sup> using Supplementary equations (14), (17) and (19). The surface area per adsorption site  $S_0 = 1.12 \text{ nm}^2$

from the structure of  $\beta$ -haematin crystals<sup>10</sup>. **a**, The correlation between the velocity of a step with radius of curvature  $R$ ,  $v_R$ , and the concentrations of a step pinner (CQ or QN),  $c_p$ , and kink blocker (MQ or AQ),  $c_b$ , for the four listed inhibitor combinations. **b**, The step velocity  $v$  in the presence step pinners and kink blockers, relative to that in pure solutions  $v_0$ , at the listed constant ratios of kink blocker to step pinner, corresponding to experimental determinations in Fig. 2c, e, compared with  $v$  in the presence of the listed step pinners only.



**Extended Data Fig. 6 | The regions of antagonistic and synergistic cooperativity in the plane of the concentrations of step pinners  $c_p$  and kink blockers  $c_B$ .** Solid line represents the equation  $(\partial v_R / \partial c_B)_{c_H, c_P} = 0$ , where  $(\partial v_R / \partial c_B)_{c_H, c_P}$  follows Supplementary equation (28). This line corresponds to additive cooperativity and divides the  $(c_p, c_B)$  plane into fields where  $(\partial v_R / \partial c_B)_{c_H, c_P} < 0$  marks that step pinners and kink blockers cooperate synergistically, and  $(\partial v_R / \partial c_B)_{c_H, c_P} > 0$  indicates antagonistic cooperativity between the two inhibitors.

**Extended Data Table 1 | The combination index for the four listed step pinner/kink blocker pairs**

IC	CQ/MQ	CQ/AQ	QN/MQ	QN/AQ		
Crystal Length						
10	0.223	1.628	0.232	0.359		
20	0.217	2.548	0.230	0.419		
30	0.215	2.545	0.227	0.460		
40	0.698	1.560	0.297	0.232		
50	1.089	1.266	0.318	0.240		
Step Velocity						
10	1.260	1.263	2.200	0.875		
20	1.251	1.917	2.256	1.107		
30	1.164	1.783	1.984	0.968		
40	1.208	1.325	1.962	0.920		
50	1.039	1.099		0.975		
Nucleation Rate						
10	0.460	0.372	0.438	0.470		
20	0.995	1.033	0.726	0.843		
30	1.403	1.608	1.072	1.153		
40	1.603	1.983	1.306	1.285		
50	1.731	2.189	1.525	1.394		
≤0.7	0.7 - 0.85	0.85 - 0.9	0.9 - 1.1	1.1 - 1.2	1.2 - 1.3	≥1.3
Synergism	Moderate Synergism	Slight Synergism	Nearly Additive	Slight Antagonism	Moderate Antagonism	Antagonism

The combination index was calculated for the inhibition of crystal length in bulk crystallization experiments (corresponding to isobolograms in Extended Data Fig. 1d), the step velocity (isobolograms in Extended Data Fig. 2a) and the two-dimensional nucleation rate of new crystal layers (isobolograms in Extended Data Fig. 2b). A classification of combination index values as synergy, additivity and antagonism is provided at the bottom.



Extended Data Table 2 | The analysis of variance test parameters

Source of Variation	Sum of Squares	Degrees of freedom	Mean squares	F	P-value	F critical
No inhibitor/AQ						
Between Groups	100.7	1	100.7	9.43	0.00247	3.89
Within Groups	1858.3	194	10.6			
Total	1959.1	195			Confidence interval	95%
No inhibitor/MQ						
Between Groups	327.2	1	327.2	23.15	3.50E-06	3.90
Within Groups	2204.9	176	14.1			
Total	2532.2	177			Confidence interval	95%
MQ/AQ						
Between Groups	306.7	1	306.7	25.3	8.33E-07	3.87
Within Groups	3581.9	296	12.1			
Total	3888.6	297			Confidence interval	95%

The parameters were used to test the distinction between the values of the surface free energy  $\gamma$  in haematin solution in the absence of inhibitors and in the presence of AQ or MQ.

**Extended Data Table 3 | Concentrations of free haematin [H], free inhibitors [D] and kink blocker–haematin complexes [H<sub>2</sub>B]**

<b>c<sub>H</sub>, mM</b>	<b>C<sub>B</sub>, μM</b>	<b>[H], mM</b>	<b>[B], μM</b>	<b>[H<sub>2</sub>B], μM</b>	<b>[B]<sup>-1</sup>, μM<sup>-1</sup></b>	<b>[H<sub>2</sub>B]<sup>-1</sup>, μM<sup>-1</sup></b>	<b>v'<sub>0</sub>v<sub>0</sub><sup>-1</sup></b>	<b>v, nm s<sup>-1</sup></b>	<b>v'<sub>0</sub>(v'<sub>0</sub> – v)<sup>-1</sup></b>
<b>MQ</b>									
0.28	2.0	0.278	0.948	1.025	1.054	0.976	0.983	0.654	2.990
0.28	3.0	0.277	1.486	1.594	0.673	0.627	0.973	0.673	3.239
0.28	4.0	0.276	1.936	2.063	0.516	0.485	0.966	0.579	2.500
0.28	9.0	0.271	4.501	4.619	0.222	0.216	0.923	0.500	2.182
0.28	14.0	0.266	7.106	7.034	0.140	0.142	0.883	0.470	2.1403
<b>AQ</b>									
0.28	0.66	0.279	0.016	0.65	60.5	1.534	0.989	0.776	4.637
0.28	1.2	0.278	0.029	1.15	34.0	0.864	0.981	0.711	3.632
0.28	1.5	0.277	0.038	1.48	26.0	0.672	0.975	0.679	3.298
0.28	2.0	0.276	0.051	2.00	19.3	0.498	0.967	0.623	2.814
0.28	4.1	0.272	0.106	4.01	9.41	0.249	0.934	0.583	2.663
0.28	7.1	0.266	0.190	6.95	5.21	0.144	0.885	0.584	2.948
0.28	15.0	0.251	0.457	14.6	2.19	0.068	0.755	0.635	6.302

Concentrations modified by inhibitor–haematin complexation were evaluated at analytical concentrations of haematin c<sub>H</sub> and inhibitor C<sub>B</sub> using complexation constants 14 and 510 mM<sup>-2</sup> for MQ and AQ, respectively<sup>13</sup>. Evaluation of v'<sub>0</sub> lowered from the step velocity in the absence of inhibitors v<sub>0</sub> owing to the decrease of haematin concentration from c<sub>H</sub> to [H]. The variables v'<sub>0</sub>(v'<sub>0</sub> – v)<sup>-1</sup> and c<sub>B</sub><sup>-1</sup> of the linearized form of the correlation between v and c<sub>B</sub>, Supplementary equation (7), for c<sub>B</sub> = [B] and c<sub>B</sub> = [H<sub>2</sub>B], respectively.

Extended Data Table 4 | The Langmuir constant for adsorption of MQ and AQ at kinks

	Assuming unliganded MQ and AQ are the active inhibitors		Assuming the complexes H <sub>2</sub> MQ and H <sub>2</sub> AQ are the active inhibitors	
	MQ	AQ	MQ	AQ
$K_{LB}, \mu\text{M}^{-1}$	1.9	50	1.64	1.74
$\xi$	0.54	0.4	0.54	0.4
$[B]$ or $[H_2B]$ at $C_B = 2.5 \mu\text{M}$	1.2	0.06	1.3	2.45
$\xi K_{LB} C_B$	1.22	1.09	1.15	1.7
$\xi \ln(1 + K_{LB} C_B)$	0.64	0.55	0.62	0.67

The Langmuir constant  $K_{LB}$  and the limiting fraction of occupied kinks  $\xi$  determined from the linear plots in Extended Data Fig. 4 assuming that unliganded MQ and AQ are the active inhibitors and, alternatively, that the complexes  $H_2MQ$  and  $H_2AQ$  are the active inhibitors. Evaluation of  $[MQ]$  and  $[AQ]$  or  $[H_2MQ]$  and  $[H_2AQ]$  at  $c_b = 2.5 \mu\text{M}$ , at which the inhibitor effects on the surface free energy of the step edge  $\gamma$  were measured, and of the factors  $\xi K_{LB}[B]$ ,  $\xi \ln(1 + K_{LB}[B])$ ,  $\xi K_{LB}[H_2B]$  and  $\xi \ln(1 + K_{LB}[H_2B])$  used in the evaluation of  $\Delta\gamma$  in the presence of an inhibitor.

# Superstructure control of first-cycle voltage hysteresis in oxygen-redox cathodes

<https://doi.org/10.1038/s41586-019-1854-3>

Received: 12 February 2019

Accepted: 1 October 2019

Published online: 9 December 2019

Robert A. House<sup>1</sup>, Urmimala Maitra<sup>1,8</sup>, Miguel A. Pérez-Osorio<sup>1,8</sup>, Juan G. Lozano<sup>1,2</sup>, Liyu Jin<sup>1</sup>, James W. Somerville<sup>1</sup>, Laurent C. Duda<sup>3</sup>, Abhishek Nag<sup>4</sup>, Andrew Walters<sup>4</sup>, Ke-Jin Zhou<sup>4</sup>, Matthew R. Roberts<sup>1</sup> & Peter G. Bruce<sup>1,5,6,7\*</sup>

In conventional intercalation cathodes, alkali metal ions can move in and out of a layered material with the charge being compensated for by reversible reduction and oxidation of the transition metal ions. If the cathode material used in a lithium-ion or sodium-ion battery is alkali-rich, this can increase the battery's energy density by storing charge on the oxide and the transition metal ions, rather than on the transition metal alone<sup>1–10</sup>. There is a high voltage associated with oxidation of O<sup>2–</sup> during the first charge, but this is not recovered on discharge, resulting in reduced energy density<sup>11</sup>. Displacement of transition metal ions into the alkali metal layers has been proposed to explain the first-cycle voltage loss (hysteresis)<sup>9,12–16</sup>. By comparing two closely related intercalation cathodes, Na<sub>0.75</sub>[Li<sub>0.25</sub>Mn<sub>0.75</sub>]O<sub>2</sub> and Na<sub>0.6</sub>[Li<sub>0.2</sub>Mn<sub>0.8</sub>]O<sub>2</sub>, here we show that the first-cycle voltage hysteresis is determined by the superstructure in the cathode, specifically the local ordering of lithium and transition metal ions in the transition metal layers. The honeycomb superstructure of Na<sub>0.75</sub>[Li<sub>0.25</sub>Mn<sub>0.75</sub>]O<sub>2</sub>, present in almost all oxygen-redox compounds, is lost on charging, driven in part by formation of molecular O<sub>2</sub> inside the solid. The O<sub>2</sub> molecules are cleaved on discharge, reforming O<sup>2–</sup>, but the manganese ions have migrated within the plane, changing the coordination around O<sup>2–</sup> and lowering the voltage on discharge. The ribbon superstructure in Na<sub>0.6</sub>[Li<sub>0.2</sub>Mn<sub>0.8</sub>]O<sub>2</sub> inhibits manganese disorder and hence O<sub>2</sub> formation, suppressing hysteresis and promoting stable electron holes on O<sup>2–</sup> that are revealed by X-ray absorption spectroscopy. The results show that voltage hysteresis can be avoided in oxygen-redox cathodes by forming materials with a ribbon superstructure in the transition metal layers that suppresses migration of the transition metal.

During the first charge–discharge cycle, the cathode material Na<sub>0.75</sub>[Li<sub>0.25</sub>Mn<sub>0.75</sub>]O<sub>2</sub> exhibits voltage loss, in clear contrast to Na<sub>0.6</sub>[Li<sub>0.2</sub>Mn<sub>0.8</sub>]O<sub>2</sub> which does not, despite their very similar compositions (Fig. 1a, c). Both materials possess the P2-type structure (Extended Data Fig. 1), composed of Na<sup>+</sup> ions in trigonal prismatic (P) coordination and with two transition metal (TM) oxide, TMO<sub>2</sub>, slabs required to describe the repeat stacking sequence (Fig. 1b). However, they exhibit different superstructures—specifically, different ordering of the Li and Mn in the TM layer (Fig. 1d, e). Na<sub>0.75</sub>[Li<sub>0.25</sub>Mn<sub>0.75</sub>]O<sub>2</sub> has honeycomb ordering, as observed in the majority of O-redox materials, whereas Na<sub>0.6</sub>[Li<sub>0.2</sub>Mn<sub>0.8</sub>]O<sub>2</sub> has a different ordering, composed of ribbons of Mn (Extended Data Fig. 2).

Confirmation that both materials are dominated by O-redox was obtained by operando electrochemical mass spectrometry (OEMS) and Mn L-edge X-ray absorption spectroscopy (XAS) along with resonant inelastic X-ray scattering (RIXS). In the case of Na<sub>0.75</sub>[Li<sub>0.25</sub>Mn<sub>0.75</sub>]O<sub>2</sub>, the data demonstrating that this is an O-redox compound are reported

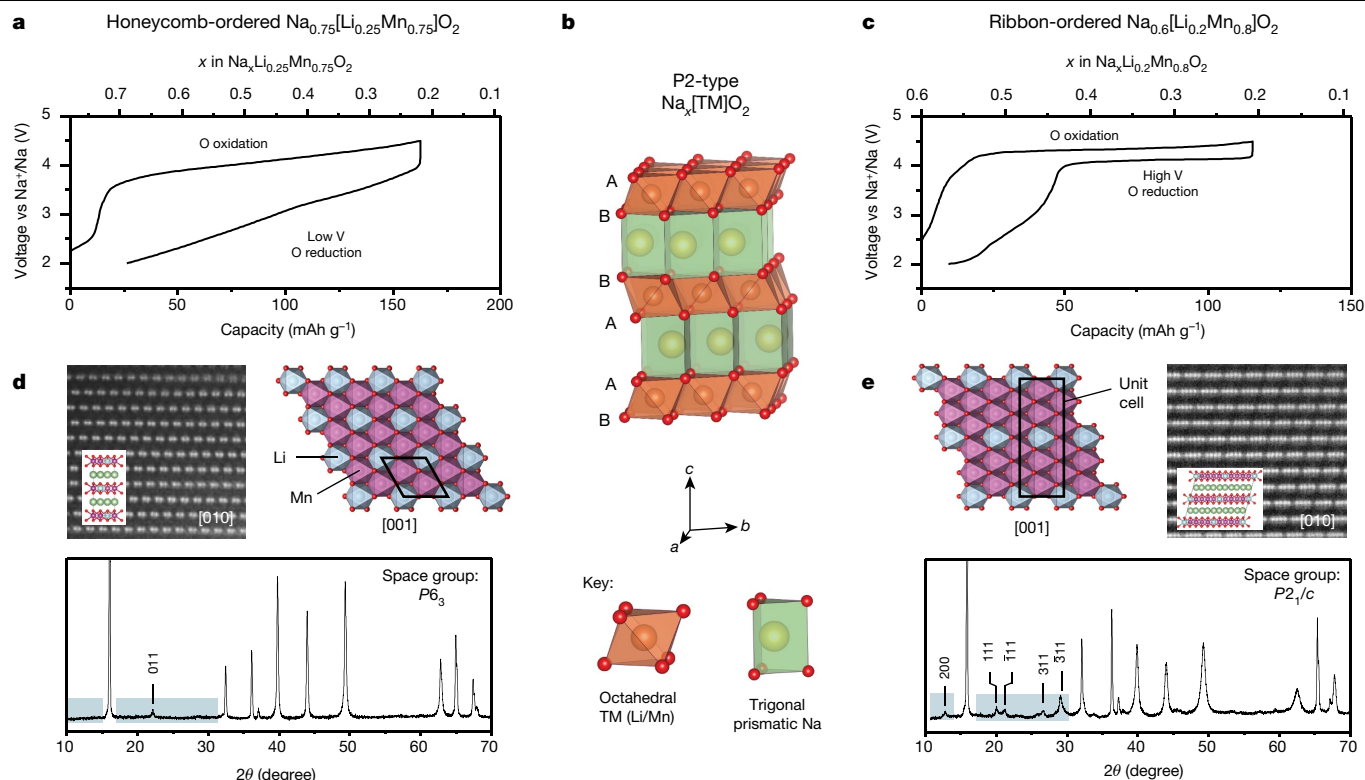
elsewhere; no O-loss was observed<sup>17</sup>. Similarly, for Na<sub>0.6</sub>[Li<sub>0.2</sub>Mn<sub>0.8</sub>]O<sub>2</sub>, no evidence of O-loss is observed as seen by OEMS (Extended Data Fig. 3). XAS and RIXS identifies electron holes on O (Extended Data Fig. 4).

## Honeycomb superstructure lost, ribbon retained

Powder X-ray diffraction (PXRD) data for Na<sub>0.6</sub>[Li<sub>0.2</sub>Mn<sub>0.8</sub>]O<sub>2</sub> are presented in Fig. 2b. At the end of charge, the diffraction peaks belonging to the P2 phase have reduced in intensity. New peaks, notably the broad peak at 16.5° in 2θ and peaks at 37° and 66°, have appeared. These peaks correspond to the most prominent peaks indexed on an O2 structure (002, 101 and 110 peaks, respectively). Similar changes in the PXRD have been observed for other charged P2-type Na[TM]O<sub>2</sub> compounds<sup>18–22</sup>. Upon sufficient desodiation, the TMO<sub>2</sub> slabs glide along a number of unique crystallographic vectors, changing the coordination environment of ions in the alkali metal (AM) layers from trigonal prismatic (P) to octahedral (O), with reduced interlayer spacing. These phases

<sup>1</sup>Department of Materials, University of Oxford, Oxford, UK. <sup>2</sup>Departamento de Ingeniería y Ciencia de los Materiales y del Transporte, Escuela Técnica Superior de Ingeniería, Universidad de Sevilla, Sevilla, Spain. <sup>3</sup>Department of Physics and Astronomy, Division of Molecular and Condensed Matter Physics, Uppsala University, Uppsala, Sweden. <sup>4</sup>Diamond Light Source, Harwell, UK. <sup>5</sup>Department of Chemistry, University of Oxford, Oxford, UK. <sup>6</sup>The Henry Royce Institute, Oxford, UK. <sup>7</sup>The Faraday Institution, Didcot, UK. <sup>8</sup>These authors contributed equally: Urmimala Maitra, Miguel A. Pérez-Osorio. \*e-mail: peter.bruce@materials.ox.ac.uk





**Fig. 1 | Electrochemistry and structure of honeycomb- and ribbon-ordered cathode materials.** **a, c**, First-cycle voltage curves for (a)  $\text{Na}_{0.75}[\text{Li}_{0.25}\text{Mn}_{0.75}]\text{O}_2$  and (c)  $\text{Na}_{0.6}[\text{Li}_{0.2}\text{Mn}_{0.8}]\text{O}_2$ . As discussed in the text, the dominant species extracted on charge is  $\text{Na}^+$ , not  $\text{Li}^+$ . Discharge limit, 2.0 V; charge limit, 4.5 V; rate  $10 \text{ mA g}^{-1}$ . **b**, Structural model of P2-type  $\text{Na}[\text{TM}]\text{O}_2$  with no in-plane ordering; space group  $P6_3/mmc$ . Oxide layers stack in ABBA sequence giving AM layers of trigonal prismatic  $\text{Na}^+$  sandwiched between TM layers of octahedral Mn partially substituted with Li. **d**,  $\text{Na}_{0.75}[\text{Li}_{0.25}\text{Mn}_{0.75}]\text{O}_2$  possesses the P2 structure with the well-known honeycomb superstructure ordering of Li and Mn within the TM layer; space group  $P6_3$ . Two-atom dumbbells (Mn–Mn) along the [010] direction, which are characteristic of the honeycomb superstructure, are seen

with ADF-STEM. Note that the honeycomb superstructure dominates although the composition is not that of the ideal honeycomb ( $\text{Li}_{1/3}\text{Mn}_{2/3}$ ). Colour-coding in insets is the same as the other structural figures in **d** and **e**: purple, Mn; blue, Li; red, O, green, Na. **e**,  $\text{Na}_{0.6}[\text{Li}_{0.2}\text{Mn}_{0.8}]\text{O}_2$  possesses the same P2 structure but with the superstructure ordering of Li and Mn within the TM layer instead forming ribbons. This superstructure ordering gives rise to unique diffraction peaks in the PXRD pattern (highlighted blue) and can be fully indexed to the  $P2_1/c$  space group; see Extended Data Fig. 2. It also gives rise to four-atom dumbbells (Mn–Mn–Mn–Mn) when viewed along the [010] direction, as observed in the ADF-STEM images.

commonly exhibit broad diffraction peaks due to the existence of stacking faults<sup>18,21,22</sup>. The two-phase P2 to O2 transition is consistent with the plateau in the electrochemistry observed in Fig. 1c. The changes in the PXRD observed on charging are reversed on discharge, with the crystalline P2-phase being recovered at the end of discharge.

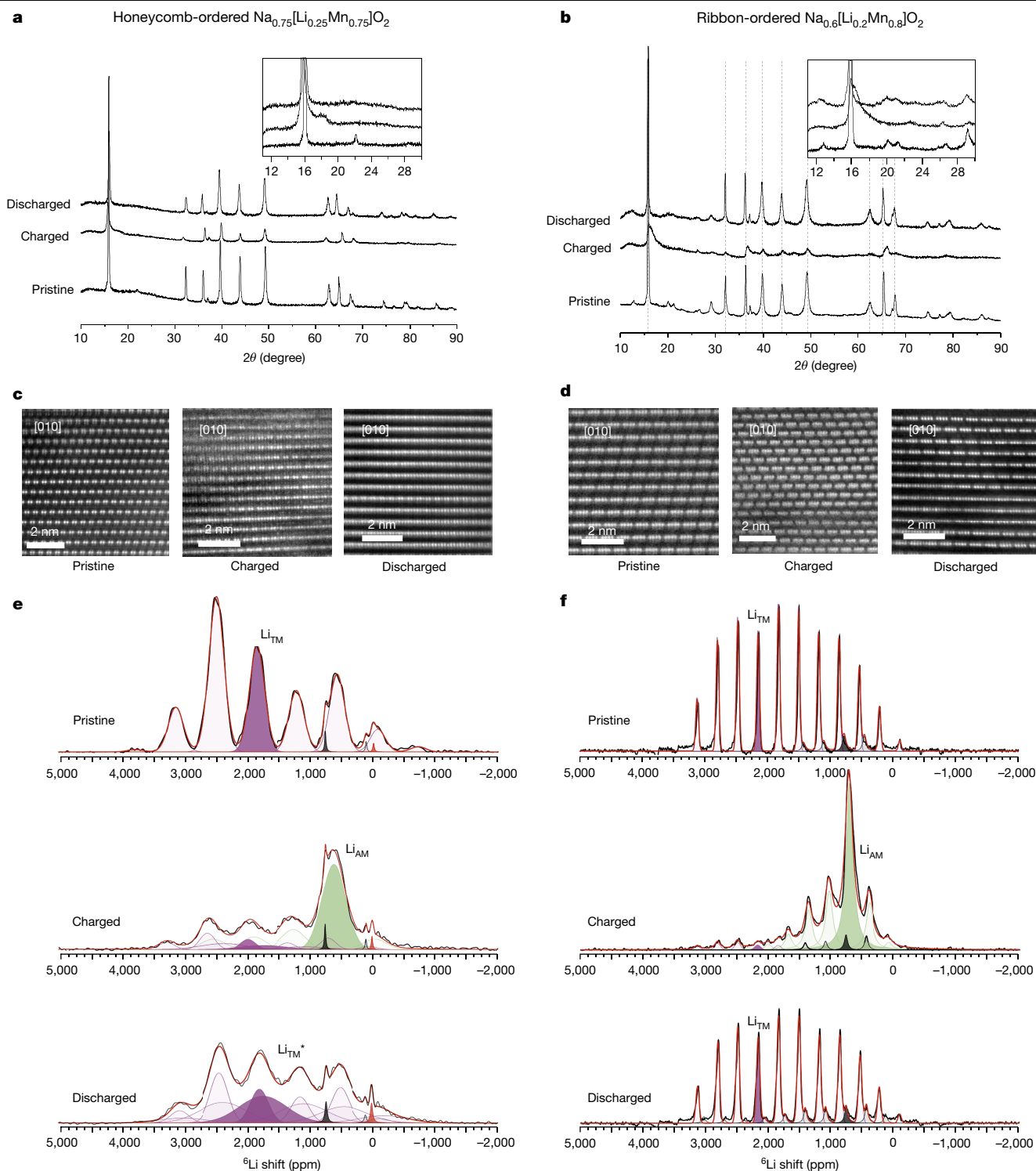
A reduction in peak intensity and peak broadening is also observed in  $\text{Na}_{0.75}[\text{Li}_{0.25}\text{Mn}_{0.75}]\text{O}_2$  upon charging (Fig. 2a), as is evidence of a new broad peak at around  $18^\circ$  characteristic of O-type layer stacking with a contracted layer spacing. It is clear, however, that the O-type phase here is not well crystallized as it does not exhibit sharp, well-defined PXRD peaks. The charging plateau for  $\text{Na}_{0.75}[\text{Li}_{0.25}\text{Mn}_{0.75}]\text{O}_2$  exhibits a gentle slope, consistent with the P2–O2 transition in this case occurring through a continuously evolving intergrowth of O stacking faults in the P structure.

Nuclear magnetic resonance data ( $^6\text{Li}$  NMR, discussed below) are sensitive to all of the Li whether in crystalline or non-crystalline regions, and thus NMR is the best technique to follow changes in Li. For both materials, NMR revealed substantial displacement of  $\text{Li}^+$  from the TM layer into sites of octahedral coordination in the AM layer upon desodiation, further confirming the presence of O-type stacking faults to accommodate the  $\text{Li}^+$  ions. We therefore conclude that transformation from P2 to O-type stacking is near-complete in both materials at high states of charge. The PXRD, being sensitive to crystalline regions, does not show so clearly the evolution of O2, especially in the case

of  $\text{Na}_{0.75}[\text{Li}_{0.25}\text{Mn}_{0.75}]\text{O}_2$ , because the O2 phase is disordered, giving rise to fewer diffraction peaks than  $\text{Na}_{0.6}[\text{Li}_{0.2}\text{Mn}_{0.8}]\text{O}_2$ . Interestingly, the diffraction peaks arising from in-plane ordering, as highlighted in Fig. 1, are recovered on discharge for  $\text{Na}_{0.6}[\text{Li}_{0.2}\text{Mn}_{0.8}]\text{O}_2$  but not for  $\text{Na}_{0.75}[\text{Li}_{0.25}\text{Mn}_{0.75}]\text{O}_2$ .

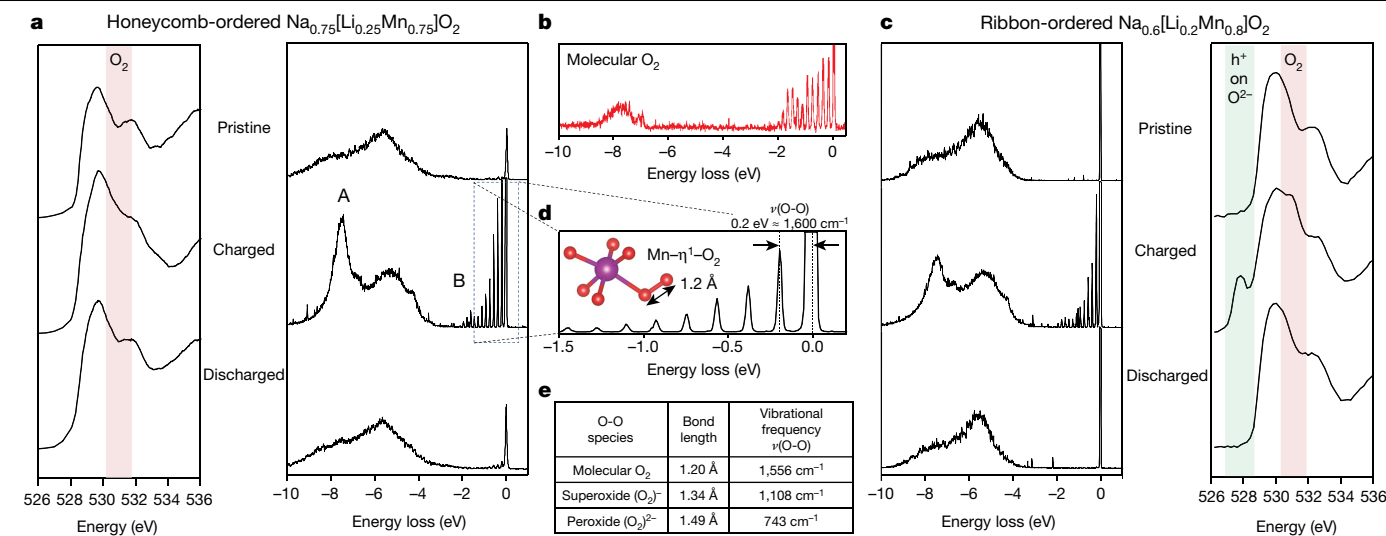
At the end of charge, the sharp line observed in the  $^6\text{Li}$  magic-angle-spinning (MAS) NMR spectrum for Li in the TM layer of pristine  $\text{Na}_{0.6}[\text{Li}_{0.2}\text{Mn}_{0.8}]\text{O}_2$  disappears, and instead a new  $\text{Li}^+$  environment with different frequency (centred at 720 ppm, shaded green) appears in its place (Fig. 2f, middle panel). This new shift is in close alignment with those observed for  $\text{Li}^+$  in octahedral coordination within AM layers in other layered compounds<sup>23</sup>. It is accordingly assigned to Li residing in the AM layer. After discharge (Fig. 2f, lower panel), the new shift disappears and the original shift is reform with the same, sharp, line shape, indicating the presence of Li again dominantly in their original octahedral TM layer sites. Further NMR data were collected at intermediate states of charge (Extended Data Fig. 5), which confirm that this process occurs via a two-phase mechanism.

The  $^6\text{Li}$  NMR results for  $\text{Na}_{0.75}[\text{Li}_{0.25}\text{Mn}_{0.75}]\text{O}_2$  also show that  $\text{Li}^+$  is displaced from the TM layers to octahedral sites in the AM layers on charge; an ensemble of shifts centred at 600 ppm is observed in the spectrum for  $\text{Na}_{0.75}[\text{Li}_{0.25}\text{Mn}_{0.75}]\text{O}_2$  on charge (Fig. 2e, middle panel). However, on discharge, as the lithium returns to vacancies in the TM layer, a considerable broadening of the ensemble of isotropic chemical shifts centred



**Fig. 2 | Evidence for the loss of honeycomb ordering and retention of ribbon ordering on the first cycle.** **a, b**, PXRD data for honeycomb-ordered P2- $\text{Na}_{0.75}[\text{Li}_{0.25}\text{Mn}_{0.75}]\text{O}_2$  and ribbon-ordered P2- $\text{Na}_{0.6}[\text{Li}_{0.2}\text{Mn}_{0.8}]\text{O}_2$ , respectively: pristine; after charging to 4.5 V; and after discharging to 2 V. Insets highlight the superstructure peak region. Superstructure peaks corresponding to the ribbon ordering reappear on discharge, indicating retention of in-plane ordering, whereas for the honeycomb the superstructure peak at  $22^\circ$  is lost irreversibly. **c, d**, STEM-ADF images. Projections are parallel to the  $ab$  plane along the [010] zone-axis. Light and dark contrast corresponds to the heavier (Mn) and lighter (Li, Na and O) scattering respectively. The ribbon superstructure is retained in the case of  $\text{Na}_{0.6}[\text{Li}_{0.2}\text{Mn}_{0.8}]\text{O}_2$  on charge and discharge, although there is some evidence of disorder in the discharged state, which is discussed in the text. In contrast, the two-atom (Mn–Mn) dumbbells

within the TM layers almost completely disappear, showing that honeycomb ordering is lost. There is some evidence of scattering between the TM layers in the charged honeycomb structure, which might indicate some displacement of Mn into the AM layers, but the dominant Mn disorder is in-plane. **e, f**,  $^6\text{Li}$  MAS NMR data. Peaks corresponding to the isotropic shifts for Li in the TM layers,  $\text{Li}_{\text{TM}}$ , are shown in purple, and those corresponding to Li in AM layers,  $\text{Li}_{\text{AM}}$ , are shown in green. All other peaks are spinning sidebands. For ribbon-ordered  $\text{Na}_{0.6}[\text{Li}_{0.2}\text{Mn}_{0.8}]\text{O}_2$ , lithium migrates from TM to AM layers on charging, and this is reversed on discharge. For honeycomb-ordered  $\text{Na}_{0.75}[\text{Li}_{0.25}\text{Mn}_{0.75}]\text{O}_2$ , lithium again migrates from TM to AM layers on charging; however, on discharge, Li repopulates a TM layer with different local ordering as a result of Mn migration, labelled  $\text{Li}_{\text{TM}}^*$ . Small amounts of  $\text{Li}_2\text{MnO}_3$  and Li-containing diamagnetic impurities are shaded in black and red, respectively.



**Fig. 3 | Spectroscopic evidence for  $\text{O}_2$  formation and stable electron holes on  $\text{O}^{2-}$ .** **a, c,** Oxygen K-edge XAS and high-resolution RIXS spectra recorded at an excitation energy of 531 eV for (a) honeycomb-ordered  $\text{Na}_{0.75}[\text{Li}_{0.25}\text{Mn}_{0.75}]\text{O}_2$  and (c) ribbon-ordered  $\text{Na}_{0.6}[\text{Li}_{0.2}\text{Mn}_{0.8}]\text{O}_2$  in the pristine, charged (4.5 V), and discharged (2 V) states. The red highlighted pre-edge feature at 531 eV and RIXS features A and B are characteristic of O-redox. **b,** The high-resolution RIXS spectrum for molecular  $\text{O}_2$  at 530.3 eV (reproduced with permission from Arhammar et al.<sup>28</sup>). **d,** With high-resolution RIXS, feature B in **a** is resolved into a

progression of energy-loss peaks, arising from the vibrations of the O–O bond with a fundamental vibrational frequency,  $\nu$ , of approximately 1,600  $\text{cm}^{-1}$  matching that of molecular  $\text{O}_2$  and that expected from the 1.2-Å O–O bond in the  $\text{Mn}-\eta^1-\text{O}_2$  species predicted from DFT. **e,** Literature values for the bond lengths and frequencies of O–O dimers for comparison<sup>35</sup>. The O K-edge XAS spectrum for ribbon-ordered  $\text{Na}_{0.6}[\text{Li}_{0.2}\text{Mn}_{0.8}]\text{O}_2$  in the charged state shows the formation of stable electron holes ( $h^+$ ) on  $\text{O}^{2-}$  (green) at low energy (high voltage).

around 1,800 ppm is observed (Fig. 2e, lower panel). This is indicative of a range of new Li environments generated from different local ordering of Mn. The results are in contrast to the ribbon superstructure ordering of  $\text{Na}_{0.6}[\text{Li}_{0.2}\text{Mn}_{0.8}]\text{O}_2$ , which does not change and therefore provides the same sites for  $\text{Li}^+$  as were present in the pristine material: that is, TM ordering is largely retained in  $\text{Na}_{0.6}[\text{Li}_{0.2}\text{Mn}_{0.8}]\text{O}_2$  but not in  $\text{Na}_{0.75}[\text{Li}_{0.25}\text{Mn}_{0.75}]\text{O}_2$ . This is evidence that Mn is mobile in honeycomb-ordered  $\text{Na}_{0.75}[\text{Li}_{0.25}\text{Mn}_{0.75}]\text{O}_2$ . Integration of the Li signals for both materials shows that any loss of Li from the structure is <1% (that is, below the limit of detection) and all of the Li that is displaced returns to the TM layers. This also confirms that  $\text{Na}^+$ , not  $\text{Li}^+$ , is the dominant species removed and reinserted into the structure on charge/discharge.

Annular dark-field scanning transmission electron microscopy (ADF-STEM) data indicate that the honeycomb superstructure of the pristine  $\text{Na}_{0.75}[\text{Li}_{0.25}\text{Mn}_{0.75}]\text{O}_2$  is almost entirely lost after charging to 4.5 V. Viewed along the [010] direction (Fig. 2c), the two-atom (Mn–Mn) dumbbells are less clearly resolved in most parts of the image after charging to 4.5 V, indicating loss of the in-plane order. After discharge to 2 V, there is virtually no periodic variation in contrast along the layers, showing the honeycomb is completely lost. In contrast to this, the ADF-STEM image of  $\text{Na}_{0.6}[\text{Li}_{0.2}\text{Mn}_{0.8}]\text{O}_2$  along the [010] direction shows retention of the four-atom (Mn–Mn–Mn–Mn) configuration associated with the ribbon ordering described in Fig. 1e, with some slight disorder evident at the end of discharge. These results are in agreement with the PXRD and NMR data, which show predominantly O-type stacking in both cases. Further images from different regions of each cycled sample are included in Extended Data Fig. 6, showing the structural changes more comprehensively. It is important to note that PXRD and NMR, unlike STEM, sample the whole material, so the STEM results are representative of the material as a whole; the consistency between all three techniques reinforces the interpretation of the results.

Together the NMR, PXRD and STEM data show that the honeycomb superstructure is unstable on charging. The  $\text{Li}^+$  ions are displaced to the AM layers, and irreversible in-plane migration of Mn results in a more disordered arrangement of Mn and vacancies in the TM layer in the charged state of  $\text{Na}_{0.75}[\text{Li}_{0.25}\text{Mn}_{0.75}]\text{O}_2$ . The honeycomb ordering is not

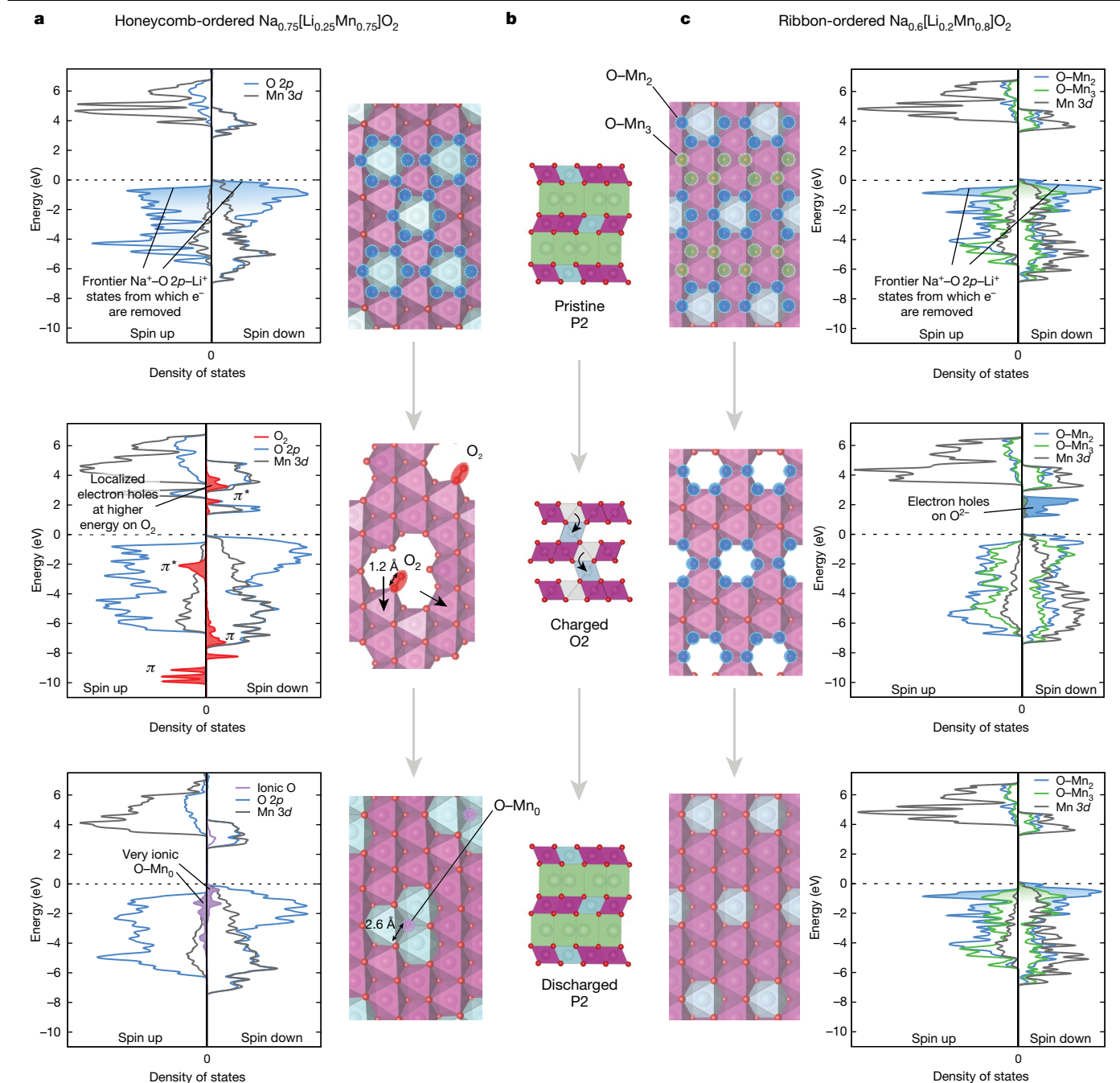
recovered on discharge, with the consequence that the  $\text{Li}^+$  ions return to different sites in the TM layer. In contrast, for the ribbon superstructure in  $\text{Na}_{0.6}[\text{Li}_{0.2}\text{Mn}_{0.8}]\text{O}_2$ , the ordering remains predominantly unchanged, and the high voltage on charge is retained on discharge.

## Molecular $\text{O}_2$ or stable $e^-$ holes

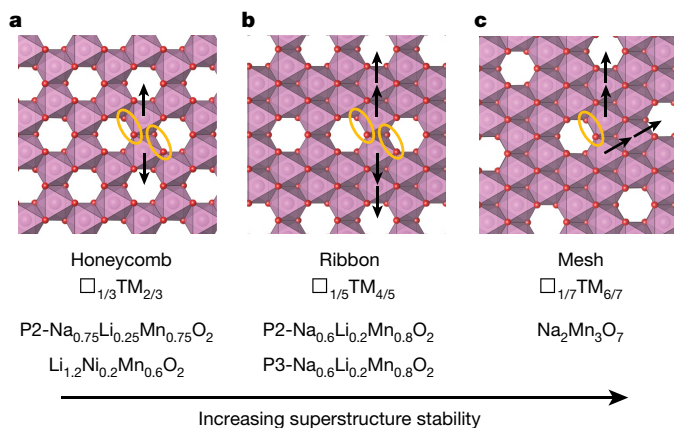
Density functional theory (DFT) calculations were performed on structural models of the charged state,  $\text{O}_2\text{-Na}_0\text{Li}_{0.25}\text{Mn}_{0.75}\text{O}_2$ , with  $\text{Li}^+$  in the AM layer, vacancies in the TM layer and different in-plane configurations of Mn. As shown in Extended Data Fig. 7a, many configurations are very similar in energy to the honeycomb arrangement (within about 20–30 meV per formula unit (f.u.), comparable to the thermal energy,  $kT = 25.7$  meV) with one notable exception that was substantially (225 meV per f.u.) lower in energy. In this arrangement, TM vacancies cluster together, resulting in replacement of the oxygens coordinated by two Mn (O–Mn<sub>2</sub>) which occurs for all O in the honeycomb structure, with O coordinated by three Mn (O–Mn<sub>3</sub>) and O atoms completely de-coordinated from Mn, which dimerize with one other O bonded to one Mn. The O–O bond length is 1.2 Å (directly comparable to that of molecular  $\text{O}_2$ , 1.208 Å), and the Mn–O distance is 2.2 Å, consistent with a weak Mn–O bond and hence formation of a  $\text{Mn}-\eta^1-\text{O}_2$  moiety (where  $\eta^1$  indicates the hapticity) containing molecular  $\text{O}_2$ . The dimerization of O to form  $\text{Mn}-\eta^1-\text{O}_2$  lowers the overall energy of the charged structure and drives the TM migration.

To evaluate what impact this structural change has on the discharge voltage, we calculated this quantity directly using the computed energies of the charged and discharged structures. To model the discharged state, the Mn and vacancy arrangement corresponding to the deep energy minimum described above was retained and the AM- and TM-layer vacancy cluster repopulated with  $\text{Na}^+$  and  $\text{Li}^+$  respectively. The resulting relaxed structure no longer possessed the short O–O distance (now 2.6 Å); the O–O bond of molecular  $\text{O}_2$  is cleaved on discharge and fully reduced  $\text{O}^{2-}$  formed (Extended Data Fig. 7c). A value of 3.2 V was obtained from the calculation of the discharge voltage, in good agreement with that observed from the electrochemistry (Fig. 1a),









**Fig. 5 | Dependence of O-redox stability on superstructure.** **a–c**, In-plane Mn migrations (arrows) required to form  $\text{O}_2$  molecules (orange ellipses) in the TM layer of charged (a) honeycomb, (b) ribbon and (c) mesh arrangements. More Mn migrations are required to form  $\text{O}_2$  in the ribbon and mesh structures, and Mn must migrate into sites already filled by Mn, making  $\text{O}_2$  formation less likely. TM layer vacancies are represented by  $\square$ .

a progression of energy-loss peaks associated with the vibrations of the O–O bond and with a fundamental vibrational frequency of about  $1,600\text{ cm}^{-1}$  (Fig. 3d), closely matching that of molecular  $\text{O}_2$  (Fig. 3b). Together with the broad inelastic peak, A, these features bear strong resemblance to the RIXS spectrum for gaseous molecular  $\text{O}_2$  (Fig. 3b)<sup>28</sup>. The ultra-high-vacuum conditions under which RIXS measurements are made ensures that the electrode materials are fully ‘out-gassed’, and hence  $\text{O}_2$  molecules in the gas phase or surface adsorbed cannot account for the RIXS observations, supporting  $\text{O}_2$  bound with  $\eta^1$  coordination to Mn. Furthermore, both spectroscopic features A and B no longer appear in the discharged sample, indicating that the  $\text{O}_2$  species is reduced on discharge.

Oxidation of  $\text{O}^{2-}$  results in oxygen species with an electronic driving force for O dimerization. This is reflected in the observation of molecular  $\text{O}_2$  here and reports of peroxo-like ( $\text{O}_2^{n-}$ ) with O–O bond lengths of  $1.4\text{--}2.5\text{ \AA}$  in  $4d$ - and  $5d$ -based materials<sup>6,8,29,30</sup>. Here we see clear evidence that molecular  $\text{O}_2$  (O–O  $1.2\text{ \AA}$ ) forms and not  $\text{O}_2^{2-}$  (O–O  $1.5\text{ \AA}$ ) or peroxo-like species (O–O  $1.9\text{--}2.5\text{ \AA}$ ). The results show that molecular  $\text{O}_2$  can be observed in the bulk of O-redox materials and demonstrate its important role in O-redox, in particular in voltage hysteresis. This bulk  $\text{O}_2$  is trapped in the vacancy clusters and has no mechanism of diffusing through the material to the surface; however, any  $\text{O}_2$  that is formed at the surface can escape. Although no direct  $\text{O}_2$  loss is seen for either  $\text{Na}_{0.6}[\text{Li}_{0.2}\text{Mn}_{0.8}]\text{O}_2$  or  $\text{Na}_{0.75}[\text{Li}_{0.25}\text{Mn}_{0.75}]\text{O}_2$ , some  $\text{CO}_2$  is observed in both cases across the charging plateaux. It has been shown that singlet  $\text{O}_2$  is typically evolved from O-redox materials and reacts with the electrolyte, forming  $\text{CO}_2$  (refs. 31,32). If  $\text{O}_2$  release is forced to be fast, by for example stepping to a high potential, a small amount of  $\text{O}_2$  can be detected<sup>17</sup>.

A small signature from molecular  $\text{O}_2$  is also seen in the RIXS for ribbon-ordered  $\text{Na}_{0.6}[\text{Li}_{0.2}\text{Mn}_{0.8}]\text{O}_2$  at the end of charge. This is in accord with the electrochemical data and ADF-STEM images, which show that the ribbon structure of  $\text{Na}_{0.6}[\text{Li}_{0.2}\text{Mn}_{0.8}]\text{O}_2$  is not completely preserved during the first cycle, and some low-voltage capacity is seen on the first discharge. However, turning to the O K-edge XAS data presented in Fig. 3c, in addition to the  $\text{O}_2$  feature appearing at  $531\text{ eV}$ , there is also a new feature appearing before the pre-edge at  $527.5\text{ eV}$ . This feature is exactly where electron-hole states lying just above the Fermi energy would be expected to appear and hence represents electronic states on O that can be reduced at high potential (that is, O-redox without voltage hysteresis). This is evidence of true, stable electron holes on  $\text{O}^{2-}$  (that is,  $\text{O}^{n-}$  where  $n < 2$ ) and is distinct from the localized holes that form on  $\text{O}_2$ .

## Superstructure controls voltage hysteresis

Hysteresis in O-redox materials has been related to the number of electron holes formed on oxygen, with too many resulting in structural instability and voltage loss<sup>33</sup>. However, there are a number of materials that are less oxidized than ribbon  $\text{Na}_{0.6}[\text{Li}_{0.2}\text{Mn}_{0.8}]\text{O}_2$  (0.2 electron holes,  $h^+$ , per O) yet still exhibit hysteresis, such as  $\text{Na}_2\text{RuO}_3$  (0.13  $h^+$  per O) and  $\text{Li}[\text{Ni}_{1/3}\text{Li}_{1/9}\text{Mn}_{5/9}]\text{O}_2$  (0.13  $h^+$  per O)<sup>1,10</sup>. The crucial difference here is that the latter examples both possess honeycomb-ordered TM layers, like  $\text{Na}_{0.75}[\text{Li}_{0.25}\text{Mn}_{0.75}]\text{O}_2$ , whereas  $\text{Na}_{0.6}[\text{Li}_{0.2}\text{Mn}_{0.8}]\text{O}_2$  exhibits ribbon ordering. Further examination of the literature reveals a strong evidential link between superstructure ordering and voltage hysteresis which extends across P2 and P3 Na-ion compounds and Li-rich O3 structures. Honeycomb ordering is exhibited by the vast majority of layered O-redox cathodes which consistently exhibit voltage hysteresis, and the only known examples without voltage hysteresis have a different ordering scheme. P3- $\text{Na}_{0.6}[\text{Li}_{0.2}\text{Mn}_{0.8}]\text{O}_2$  has the same ribbon ordering scheme as P2- $\text{Na}_{0.6}[\text{Li}_{0.2}\text{Mn}_{0.8}]\text{O}_2$ , but with a different stacking sequence, as we show in Extended Data Fig. 8, and does not exhibit voltage hysteresis.  $\text{Na}_2\text{Mn}_3\text{O}_7$ , which has a unique in-plane ordering scheme corresponding to its  $[\square_{1/7}\text{Mn}_{6/7}]$  TM layer composition (where  $\square$  represents a vacancy), also shows no voltage hysteresis<sup>14,15,34</sup>.

The highest-energy O  $2p$  states in pristine honeycomb  $\text{Na}_{0.75}[\text{Li}_{0.25}\text{Mn}_{0.75}]\text{O}_2$  and ribbon  $\text{Na}_{0.6}[\text{Li}_{0.2}\text{Mn}_{0.8}]\text{O}_2$  are those coordinated by two ionic cations ( $\text{Li}^+$ ,  $\text{Na}^+$ ) from the TM and AM layer respectively, forming  $\text{Na}^+-\text{O } 2p-\text{Li}^+$  dumbbells<sup>4,5</sup>. On charging, electrons are removed from these states, oxidizing  $\text{O}^{2-}$  to form  $\text{O}^{n-}$  ( $n < 2$ ) and triggering displacement of  $\text{Li}^+$  from the TM to AM layers (Fig. 4b). For perfect honeycomb ordering, all oxide ions are coordinated in the TM layer by two Mn, O–Mn<sub>2</sub>, are degenerate in energy and are equally susceptible to oxidation on charging. However, this degeneracy can be broken through Mn migration which changes the Mn coordination of the  $\text{O}^{n-}$  ions from O–Mn<sub>2</sub> to: more coordinated O–Mn<sub>3</sub>, less coordinated O–Mn<sub>1</sub> and uncoordinated O (O–Mn<sub>0</sub>) (Fig. 4a). The uncoordinated O (O–Mn<sub>0</sub>) is stabilized by dimerizing with the O–Mn<sub>1</sub>, forming the Mn– $\eta^1$ - $\text{O}_2$  moieties, as demonstrated by DFT and RIXS above. It is this electronic driving force that promotes Mn migration to disrupt the honeycomb superstructure. Discharge involves reduction of the unoccupied states on  $\text{O}_2$ , triggering cleavage of the O–O bond, the formation of fully reduced  $\text{O}^{2-}$  and the return of  $\text{Li}^+$  to the TM layers. However, now they return to different sites, instead occupying sites in the vacancy cluster. The  $\text{Na}^+$  ions return to the AM layers. The discharge voltage for this process ( $3.2\text{ V}$ , as noted above) is much lower than the voltage on charge, explaining the first-cycle voltage hysteresis of the honeycomb superstructured O-redox cathode. In the case of the ribbon superstructure, Mn migration is suppressed, preventing  $\text{O}_2$  formation and stabilizing electron holes on  $\text{O}^{2-}$  (Figs. 3c, 4c). In the charged honeycomb structure, only two Mn are required to migrate into adjacent vacancies to generate free  $\text{O}^{n-}$  that can pair with a neighbouring O to form  $\text{O}_2$  (Fig. 5a). In contrast, because vacancies are more dispersed in the ribbon structure, multiple Mn displacements, including sequential Mn hops, would be required to form the TM vacancy clusters (Fig. 5b). Ribbon ordering thus provides increased stability for high-voltage O redox ( $4.1\text{ V}$  from calculation) by preserving the degeneracy of the O  $2p$  states.

Ribbon ordering is not, however, 100% stable. Even on the first cycle not all the charge capacity at  $4.3\text{ V}$  is recovered on discharge. Furthermore, Extended Data Fig. 9 shows that dwelling for increasing time in the highly desodiated charged state promotes the loss of the high-voltage discharge, suggesting increased Mn migration. Upon extended cycling, the discharge plateaux gradually decrease in length, with greater evidence of low-voltage capacity similar to that of  $\text{Na}_{0.75}[\text{Li}_{0.25}\text{Mn}_{0.75}]\text{O}_2$ . PXRD data (Extended Data Fig. 10) confirm that after 10 cycles the diffraction peaks arising from the superstructure ordering are reduced, and there is increasing strain broadening within the  $a$ – $b$  plane indicative of Mn migration and loss of the ribbon

ordering. ADF-STEM imaging (Extended Data Fig. 6) also shows loss of ribbon ordering after 10 cycles. The loss of voltage correlates with the loss of superstructure, further reinforcing the relationship between the two. This irreversibility of the high-voltage plateau is also seen for the P3-type analogue of P2- $\text{Na}_{0.6}[\text{Li}_{0.2}\text{Mn}_{0.8}]\text{O}_2$  (refs. <sup>14,15</sup>), which exhibits the same ribbon superstructure ordering.

Although ribbon ordering is not completely stable, a compound possessing an ordering scheme with even more dispersed vacancy ordering has been reported that shows higher reversibility of the high-voltage O-redox plateau (Fig. 5c)<sup>34</sup>. This observation, underpinned by our work revealing the critical role of superstructure in preserving high-voltage O-redox, defines a compelling strategy in the search for high-energy-density Li-rich cathodes.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-019-1854-3>.

- Lu, Z., Beaulieu, L. Y., Donabarger, R. A., Thomas, C. L. & Dahn, J. R. Synthesis, structure, and electrochemical behavior of  $\text{Li}[\text{Ni}_x\text{Li}_{1/3-2x/3}\text{Mn}_{2/3-x/3}]\text{O}_2$ . *J. Electrochem. Soc.* **149**, A778–A791 (2002).
- Johnson, C. S. et al. The significance of the  $\text{Li}_2\text{MnO}_3$  component in ‘composite’  $x\text{Li}_2\text{MnO}_3 \cdot (1-x)\text{LiMn}_{0.5}\text{Ni}_{0.5}\text{O}_2$  electrodes. *Electrochem. Commun.* **6**, 1085–1091 (2004).
- Koga, H. et al. Reversible oxygen participation to the redox processes revealed for  $\text{Li}_{1.20}\text{Mn}_{0.54}\text{Co}_{0.13}\text{Ni}_{0.13}\text{O}_2$ . *J. Electrochem. Soc.* **160**, A786–A792 (2013).
- Luo, K. et al. Charge-compensation in 3d-transition-metal-oxide intercalation cathodes through the generation of localized electron holes on oxygen. *Nat. Chem.* **8**, 684–691 (2016).
- Seo, D.-H. et al. The structural and chemical origin of the oxygen redox activity in layered and cation-disordered Li-excess cathode materials. *Nat. Chem.* **8**, 692–697 (2016).
- Saubanère, M., McCalla, E., Tarascon, J.-M. & Doublet, M.-L. The intriguing question of anionic redox in high-energy density cathodes for Li-ion batteries. *Energy Environ. Sci.* **9**, 984–991 (2016).
- Oishi, M. et al. Direct observation of reversible oxygen anion redox reaction in Li-rich manganese oxide,  $\text{Li}_2\text{MnO}_3$ , studied by soft X-ray absorption spectroscopy. *J. Mater. Chem. A* **4**, 9293–9302 (2016).
- Sathiyar, M. et al. Reversible anionic redox chemistry in high-capacity layered-oxide electrodes. *Nat. Mater.* **12**, 827–835 (2013).
- Hong, J. et al. Metal–oxygen decoordination stabilizes anion redox in Li-rich oxides. *Nat. Mater.* **18**, 256–265 (2019).
- Mortemard de Boisse, B. et al. Intermediate honeycomb ordering to trigger oxygen redox chemistry in layered battery electrode. *Nat. Commun.* **7**, 11397 (2016).
- Lu, Z. & Dahn, J. R. Understanding the anomalous capacity of  $\text{Li}/\text{Li}[\text{Ni}_x\text{Li}_{1/3-2x/3}\text{Mn}_{2/3-x/3}]\text{O}_2$  cells using in situ X-ray diffraction and electrochemical studies. *J. Electrochem. Soc.* **149**, A815 (2002).
- Gent, W. E. et al. Coupling between oxygen redox and cation migration explains unusual electrochemistry in lithium-rich layered oxides. *Nat. Commun.* **8**, 2091 (2017).
- Croy, J. R. et al. First-charge instabilities of layered-layered lithium-ion-battery materials. *Phys. Chem. Chem. Phys.* **17**, 24382–24391 (2015).
- Du, K. et al. Exploring reversible oxidation of oxygen in a manganese oxide. *Energy Environ. Sci.* **6**, 3–5 (2016).
- Rong, X. et al. Structure-induced reversible anionic redox activity in Na layered oxide cathode. *Joule* **2**, 125–140 (2018).
- Pearce, P. E. et al. Evidence for anionic redox activity in a tridimensional-ordered Li-rich positive electrode  $\beta\text{-Li}_2\text{IrO}_3$ . *Nat. Mater.* **16**, 580–586 (2017).
- House, R. A. et al. What triggers oxygen loss in oxygen redox cathode materials? *Chem. Mater.* **31**, 3293–3300 (2019).
- Yabuuchi, N. et al. A new electrode material for rechargeable sodium batteries: P2-type  $\text{Na}_{2/3}[\text{Mg}_{0.28}\text{Mn}_{0.72}]\text{O}_2$  with anomalously high reversible capacity. *J. Mater. Chem. A* **2**, 16851–16855 (2014).
- Maitra, U. et al. Oxygen redox chemistry without excess alkali-metal ions in  $\text{Na}_{2/3}[\text{Mg}_{0.28}\text{Mn}_{0.72}]\text{O}_2$ . *Nat. Chem.* **10**, 288–295 (2018).
- Tournadre, F. et al. On the mechanism of the  $\text{P2-Na}_{0.70}\text{CoO}_2 \rightarrow \text{O2-LiCoO}_2$  exchange reaction—Part I: proposition of a model to describe the  $\text{P2-O2}$  transition. *J. Solid State Chem.* **177**, 2790–2802 (2004).
- Lu, Z. & Dahn, J. R. In situ X-ray diffraction study of  $\text{P2 Na}_{2/3}[\text{Ni}_{1/3}\text{Mn}_{2/3}]\text{O}_2$ . *J. Electrochem. Soc.* **148**, A1225 (2001).
- Tournadre, F., Croguennec, L., Willmann, P. & Delmas, C. On the mechanism of the  $\text{P2-Na}_{0.70}\text{CoO}_2 \rightarrow \text{O2-LiCoO}_2$  exchange reaction—Part II: an in situ X-ray diffraction study. *J. Solid State Chem.* **177**, 2803–2809 (2004).
- Clément, R. J. et al. Direct evidence for high  $\text{Na}^+$  mobility and high voltage structural processes in  $\text{P2-Na}_x[\text{Li}_1\text{Ni}_x\text{Mn}_{1-x}]\text{O}_2$  ( $x, y, z \leq 1$ ) cathodes from solid-state NMR and DFT calculations. *J. Mater. Chem. A* **5**, 4129–4143 (2017).
- House, R. A. et al. Lithium manganese oxyfluoride as a new cathode material exhibiting oxygen redox. *Energy Environ. Sci.* **11**, 926–932 (2018).
- Wu, J. et al. Fingerprint oxygen redox reactions in batteries through high-efficiency mapping of resonant inelastic X-ray scattering. *Condens. Matter* **4**, 5 (2019).
- Xu, J. et al. Elucidating anionic oxygen activity in lithium-rich layered oxides. *Nat. Commun.* **9**, 947 (2018).
- Yang, W. & Devereaux, T. P. Anionic and cationic redox and interfaces in batteries: advances from soft X-ray absorption spectroscopy to resonant inelastic scattering. *J. Power Sources* **389**, 188–197 (2018).
- Arhammar, C. et al. Unveiling the complex electronic structure of amorphous metal oxides. *Proc. Natl Acad. Sci. USA* **108**, 6355–6360 (2011).
- McCalla, E. et al. Visualization of O–O peroxo-like dimers in high-capacity layered oxides for Li-ion batteries. *Science* **350**, 1516–1521 (2015).
- Xie, Y., Saubanère, M. & Doublet, M.-L. Requirements for reversible extra-capacity in Li-rich layered oxides for Li-ion batteries. *Energy Environ. Sci.* **10**, 266–274 (2017).
- Wandt, J., Freiberg, A. T. S., Ogrodnik, A. & Gasteiger, H. A. Singlet oxygen evolution from layered transition metal oxide cathode materials and its implications for lithium-ion batteries. *Mater. Today* **21**, 825–833 (2018).
- Freiberg, A. T. S., Roos, M. K., Wandt, J., de Vivie-Riedle, R. & Gasteiger, H. A. Singlet oxygen reactivity with carbonate solvents used for Li-ion battery electrolytes. *J. Phys. Chem. A* **122**, 8828–8839 (2018).
- Ben Yahia, M., Vergnet, J., Saubanère, M. & Doublet, M.-L. Unified picture of anionic redox in Li/Na-ion batteries. *Nat. Mater.* **18**, 496–502 (2019).
- Mortemard de Boisse, B. et al. Highly reversible oxygen-redox chemistry at 4.1 V in  $\text{Na}_{4/7-x}[\text{Li}_{1/3}\text{Mn}_{2/3}]\text{O}_2$  ( $\square$ : Mn vacancy). *Adv. Energy Mater.* **8**, 1800409 (2018).
- Radjenovic, P. M. & Hardwick, L. J. Evaluating chemical bonding in dioxides for the development of metal–oxygen batteries: vibrational spectroscopic trends of dioxygenyls, dioxygen, superoxides and peroxides. *Phys. Chem. Chem. Phys.* **21**, 1552–1563 (2019).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2019

## Methods

### Synthesis

$\text{Na}_{0.6}[\text{Li}_{0.2}\text{Mn}_{0.8}]\text{O}_2$  and  $\text{Na}_{0.75}[\text{Li}_{0.25}\text{Mn}_{0.75}]\text{O}_2$  were prepared via solid-state reaction between stoichiometric amounts of  $\text{Na}_2\text{CO}_3$  ( $\geq 99.0\%$ , Aldrich),  $\text{Li}_2\text{CO}_3$  ( $\geq 99.0\%$ , Aldrich) and  $\text{MnO}_2$  ( $\geq 99.0\%$ , Aldrich). The precursors were ball-milled for 1 h using a Retsch PM100, pressed into pellets and calcined at 800 °C for 12 h under flowing oxygen. Heating and cooling was conducted under a controlled rate of 10 °C min<sup>-1</sup> in both cases except for  $\text{Na}_{0.6}[\text{Li}_{0.2}\text{Mn}_{0.8}]\text{O}_2$  which was cooled at 2 °C min<sup>-1</sup>. As-prepared materials were transferred into an inert Ar atmosphere without exposure to air and stored for characterization. All subsequent procedures were carried out without exposure to air.

### Structural characterization

X-ray powder diffraction patterns were collected using a Cu source Rigaku diffractometer. Neutron powder diffraction patterns were collected at the POLARIS diffractometer at the ISIS neutron source. Powders were loaded into vanadium canisters and sealed under inert atmosphere for measurement. Reitveld profile refinements were performed using the GSAS suite of programs.

### Electrochemical characterization

Electrodes were prepared by mixing 80 wt% active material, 10 wt% Super P carbon and 10 wt% polytetrafluoroethylene binder in a mortar and pestle and rolling to form a self-standing film. Electrodes were incorporated into CR2032 coin cells with electrolyte-soaked ( $\text{NaPF}_6$  (Kishida) in propylene carbonate (99.7%, Sigma)) Whatman glass fibre separators and Na metal counter electrodes. Galvanostatic charge-discharge was carried out at a rate of 10 mA g<sup>-1</sup> using a Maccor Series 4000.

### Operando electrochemical mass spectrometry

Operando electrochemical mass spectrometry was carried out using a cell (ECC-Std from EL-CELL) with gas inlet and outlet ports. Argon carrier gas was flowed at constant rate (0.8 µl min<sup>-1</sup> Bronkhurst mass-flow controller) through the cell and into a quadrupole mass spectrometer (Thermo Fischer) equipped with turbomolecular pump (Pfeiffer Vacuum).

### Solid-state NMR

Solid-state NMR experiments were performed on a 400-MHz Bruker Avance III HD spectrometer at the <sup>6</sup>Li Larmor frequency of 58.99 MHz. All spectra were recorded with a rotor-synchronized Hahn-echo pulse sequence. The <sup>6</sup>Li spectra were externally referenced with LiCl aqueous solution at 0.0 ppm.

For  $\text{Na}_{0.6}[\text{Li}_{0.2}\text{Mn}_{0.8}]\text{O}_2$ , a 3.2-mm MAS probe was used. The MAS rate was 19 kHz, and the probe temperature was controlled at 268 K. The applied  $\pi/2$  pulse length was 3.5 µs and the delay between  $\pi/2$  and  $\pi$  pulses was 47.4 µs (one rotor period). The transmitter frequency was set to 1,600 ppm. For  $\text{Na}_{0.75}[\text{Li}_{0.25}\text{Mn}_{0.75}]\text{O}_2$ , a 1.9-mm MAS probe was used with MAS rate of 38 kHz, and the probe temperature was set to 298 K. The applied  $\pi/2$  pulse length was 2 µs and the delay between  $\pi/2$  and  $\pi$  pulses was 23.3 µs (one rotor period). The transmitter frequency was set to 1,600 ppm.

The spectra were normalized by the total number of scans and the weight of active materials packed in the rotors. Spectra fitting and deconvolution were carried out with the dmfit program<sup>36</sup>.

### ADF-STEM

ADF-STEM micrographs were collected on an aberration-corrected JEOL ARM 200F operated at 200 kV. The convergence semi-angle used was 22 mrad, and the collection semi-angle was 69.6–164.8 mrad (ADF). In all cases, sets of fast-acquisition multi-frame images were recorded and subsequently corrected for drift and scan distortions using SmartAlign<sup>37</sup>.

### Computation

DFT calculations including Hubbard corrections<sup>38</sup> were performed using Quantum Espresso<sup>39</sup>. We used the Perdew, Burke and Ernzerhof (PBE)<sup>40</sup> exchange-correlation functional. The core-valence interaction was taken into account by using the projector-augmented wave (PAW) method<sup>41</sup>. The wavefunctions were represented through a plane-wave basis set with an energy cut-off of 70 Ry. Spin polarization was included. All calculations were performed considering a ferromagnetic ordering of Mn atoms. A Hubbard U parameter of 4 eV for Mn 3d states was used, similar to that reported for other closely related compounds<sup>6,42</sup>. To find the *k*-point condition, the total energy of the supercell was converged with respect to the number of *k* points, and convergence was reached with a 2 × 2 × 2 Monkhorst–Pack *k*-point grid. Crystal structures were relaxed until forces on the atoms were less than 0.08 eV Å<sup>-1</sup> and the total stresses on the cell were less than 0.05 kbar. The supercell for  $\text{Na}_{0.75}[\text{Li}_{0.25}\text{Mn}_{0.75}]\text{O}_2$  contains 90 atoms: 18 Na atoms; 6 Li atoms; 18 Mn atoms; and 48 O atoms.

For electronic structure calculations, we carried out spin-polarized DFT calculations using the HSE functional. An exact exchange mixing parameter of 0.25 was used for all calculations. Norm-conserving pseudopotentials were used to describe the core-valence interaction<sup>43</sup>. The electronic wavefunctions were described using a plane-wave basis set with an energy cut-off of 80 Ry. A Monkhorst–Pack *k*-point grid of 2 × 2 × 2 was used. The input structures were obtained from the DFT+U lattice relaxations, and the nuclear positions were allowed to further relax at the HSE level, keeping the lattice parameters fixed.

Intercalation–deintercalation voltages (*V*) were computed using the Nernst equation,  $V = \Delta G/(zF)$ , where  $\Delta G$  is the Gibbs free energy change, *F* is the Faraday constant and *z* is the charge that is transferred. The change in the Gibbs free energy is defined as  $\Delta G = \Delta E + P\Delta V - T\Delta S$ , where *P* and *T* are pressure and temperature, respectively, and  $\Delta E$ ,  $\Delta V$  and  $\Delta S$  are the change in internal energy, volume and entropy, respectively. The first-principles calculations were carried out at 0 K and zero pressure. Under these conditions, the Gibbs free energy change is then given by the change in the internal energy,  $\Delta G = \Delta E$ . Thus, for the sodium deintercalation–intercalation reaction  $\text{Na}_{x_1}\text{TMo}_2 \rightarrow \text{Na}_{x_2}\text{TMo}_2 + (x_1 - x_2)\text{Na}$ , the voltage is given by:

$$V = - \frac{E(\text{Na}_{x_1}\text{TMo}_2) - E(\text{Na}_{x_2}\text{TMo}_2) - (x_1 - x_2)E(\text{Na})}{(x_2 - x_1)F}$$

where  $x_1 > x_2$  and  $E(\text{Na}_{x_1}\text{TMo}_2)$  and  $E(\text{Na}_{x_2}\text{TMo}_2)$  are the internal energies of the sodiated and desodiated transition metal (Tm) oxides, respectively, and  $E(\text{Na})$  is the internal energy of metallic sodium. These quantities are obtained directly from the first-principles calculations. This procedure is well established<sup>44,45</sup>.

For phases with partial Na occupancy, Na ordering was investigated using combinatorics. Simple random sampling was used to choose a representative subset of non-symmetry equivalent configurations for relaxation. A similar methodology was used to investigate Mn disorder in the charged honeycomb phase. Fully desodiated models were prepared for the charged phases to make calculations more computationally tractable.

### Spectroscopic characterization

Soft XAS and high-resolution RIXS data were recorded at i21 Diamond Light Source in the UK with supporting data from BL27SU of the RIKEN/JASRI Spring8 synchrotron in Japan and the ADDRESS beamline at the Swiss Light Source. Mn L-edge data were collected in inverse partial fluorescence yield mode, and O K-edge data are plotted in the partial fluorescence yield mode, both of which are bulk-sensitive methods.

## Data availability

Supporting research data have been deposited in the Oxford Research Archive and will be available at <https://ora.ox.ac.uk/objects/uuid:646b18a1-88b0-4575-8282-2bcdcb20a7d>.

36. Massiot, D. et al. Modelling one- and two-dimensional solid-state NMR spectra. *Magn. Reson. Chem.* **40**, 70–76 (2002).
37. Jones, L. et al. Smart Align—a new tool for robust non-rigid registration of scanning microscope data. *Adv. Struct. Chem. Imaging* **1**, 8 (2015).
38. Cococcioni, M. & de Gironcoli, S. Linear response approach to the calculation of the effective interaction parameters in the LDA + U method. *Phys. Rev. B* **71**, 035105 (2005).
39. Giannozzi, P. et al. QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials. *J. Phys. Condens. Matter* **21**, 395502 (2009).
40. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
41. Blöchl, P. E. Projector augmented-wave method. *Phys. Rev. B* **50**, 17953–17979 (1994).
42. Lee, D. H., Xu, J. & Meng, Y. S. An advanced cathode for Na-ion batteries with high rate and excellent structural stability. *Phys. Chem. Chem. Phys.* **15**, 3304 (2013).
43. Schlipf, M. & Gygi, F. Optimization algorithm for the generation of ONCV pseudopotentials. *Comput. Phys. Commun.* **196**, 36–44 (2015).
44. Aydinol, M. K., Kohan, A. F., Ceder, G., Cho, K. & Joannopoulos, J. Ab initio study of lithium intercalation in metal oxides and metal dichalcogenides. *Phys. Rev. B* **56**, 1354–1365 (1997).
45. Aydinol, M. K., Kohan, A. F. & Ceder, G. Ab initio calculation of the intercalation voltage of lithium–transition-metal oxide electrodes for rechargeable batteries. *J. Power Sources* **68**, 664–668 (1997).

**Acknowledgements** P.G.B. is indebted to the EPSRC, including the SUPERGEN programme, the Henry Royce Institute for Advanced Materials (EP/RO0661X/1, EP/S019367/1, EP/R010145/1)

and the Faraday Institution (FIRG007, FIRG008) for financial support. We thank H. Playford and R. Smith at ISIS, Harwell Campus for collecting neutron diffraction data. Support from the EPSRC (EP/K040375/1 ‘South of England Analytical Electron Microscope’) is also acknowledged. We acknowledge the use of the University of Oxford Advanced Research Computing (ARC) facility (<https://doi.org/10.5281/zenodo.22558>) in carrying out this work, and the resources provided by the Cambridge Tier-2 system operated by the University of Cambridge Research Computing Service (<http://www.hpc.cam.ac.uk>) funded by EPSRC Tier-2 capital grant EP/P020259/1, via the Advanced Materials for Alkali-ion Batteries (AMaIB) project. Synchrotron radiation experiments were performed at the i21 beamline, Diamond Light Source, Harwell, UK, with supporting data collected from the ADRESS beamline, Swiss Light Source, Villigen, Switzerland, and BL27SU, Spring8, Japan. We acknowledge technical and experimental support at the ADRESS beamline by T. Schmitt, D. McNally, X. Lu, L. Nue and M. Dantz and at the BL27SU beamline by K. Tsuruta. We thank N. Rees for help in collecting NMR spectra.

**Author contributions** R.A.H., U.M., M.R.R. and P.G.B. conceived the study. U.M. and R.A.H. carried out the materials synthesis, characterization and testing. R.A.H., U.M., J.W.S., M.R.R. and L.C.D. contributed to the measurement processing and interpretation of the spectroscopic data. M.A.P.-O. performed the DFT calculations. L.J. collected, processed and interpreted the NMR data. J.G.L. performed and interpreted the ADF-STEM measurements. R.A.H., A.N., A.W. and K.-J.Z. performed high-resolution RIXS and sXAS measurements. R.A.H. and P.G.B. wrote the manuscript with contributions and revisions from all authors.

**Competing interests** The authors declare no competing interests.

## Additional information

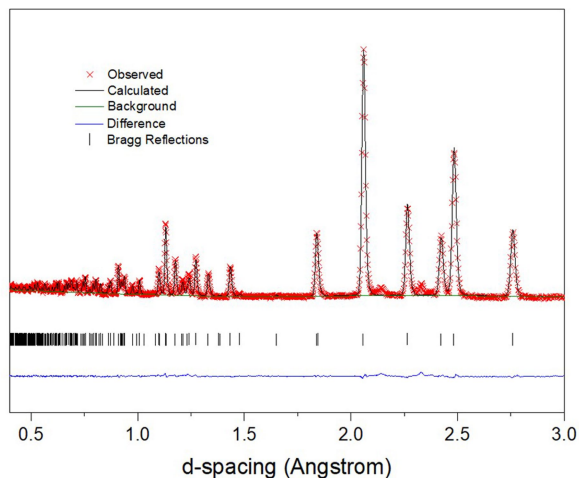
**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41586-019-1854-3>.

**Correspondence and requests for materials** should be addressed to P.G.B.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>.



**a** Honeycomb Ordered  $\text{Na}_{0.75}[\text{Li}_{0.25}\text{Mn}_{0.75}]\text{O}_2$

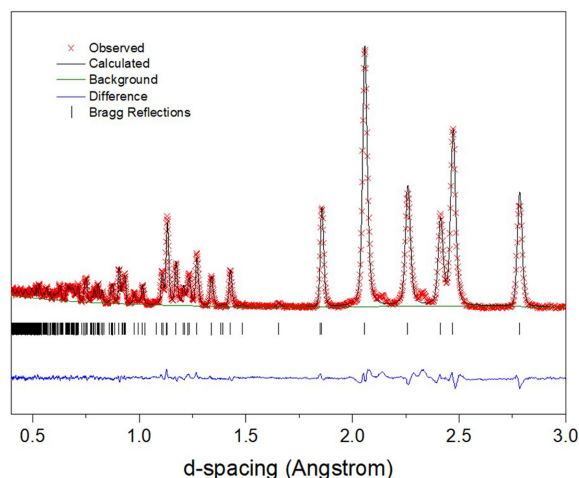


	Mult.	x	y	z	Occ.	$U_{\text{iso}}$
Li1	2	0.0	0.0	0.0	0.242(6)	0.0026(3)
Mn1	2	0.0	0.0	0.0	0.759(3)	0.0026(3)
O1	4	0.667	0.333	-0.093(1)	1.000	0.0082(1)
Na1	2	0.0	0.0	0.25	0.289(6)	0.0406(1)
Na2	2	0.667	0.333	0.25	0.486(7)	0.0406(1)

Space Group :  $P6_3/mmc$ ,  
 $a = 2.8655(1)$ ,  $c = 11.0280(2)$ ,  
G.O.F. = 1.714,  $R_w = 2.89\%$

Composition from ICP:  $\text{Na}_{0.78(2)}\text{Li}_{0.25(2)}\text{Mn}_{0.75(2)}\text{O}_2$

**b** Ribbon Ordered  $\text{Na}_{0.6}[\text{Li}_{0.2}\text{Mn}_{0.8}]\text{O}_2$



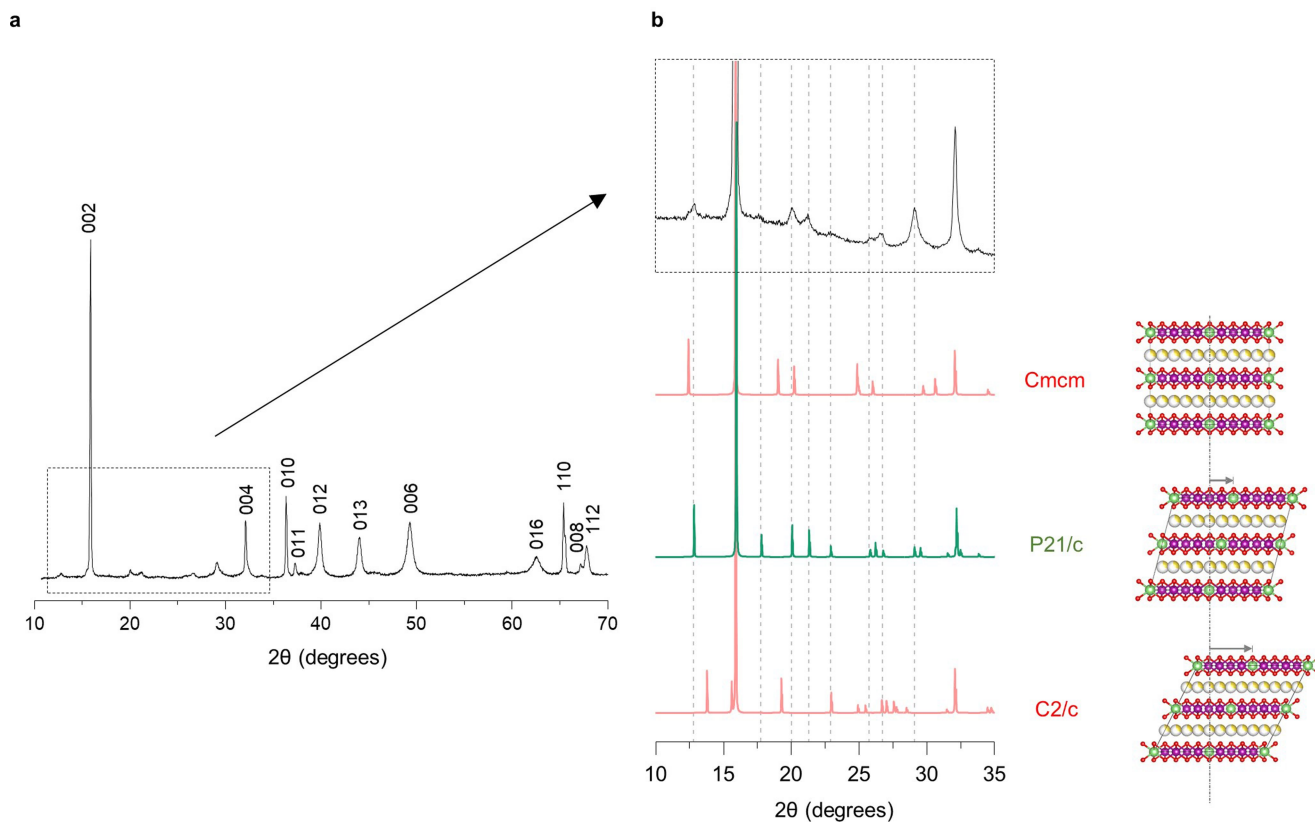
	Mult.	x	y	z	Occ.	$U_{\text{iso}}$
Li1	2	0.0	0.0	0.0	0.210(6)	0.0023(2)
Mn1	2	0.0	0.0	0.0	0.790(3)	0.0023(2)
O1	4	0.667	0.333	-0.092(1)	1.000	0.0069(1)
Na1	2	0.0	0.0	0.25	0.302(7)	0.048(2)
Na2	2	0.667	0.333	0.25	0.326(7)	0.048(2)

Space Group :  $P6_3/mmc$ ,  
 $a = 2.8513(1)$ ,  $c = 11.1301(2)$ ,  
G.O.F. = 2.305,  $R_w = 4.04\%$

Composition from ICP:  $\text{Na}_{0.63(2)}\text{Li}_{0.21(2)}\text{Mn}_{0.79(2)}\text{O}_2$

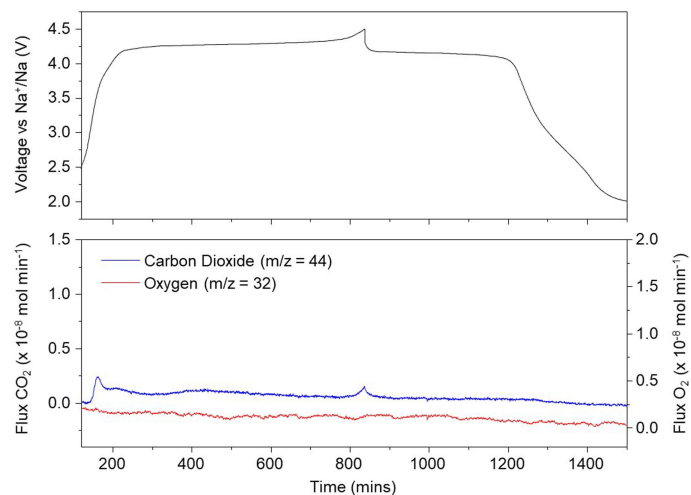
**Extended Data Fig. 1 | Further structural characterization of pristine materials.** **a, b**, Neutron powder diffraction data for **(a)**  $\text{Na}_{0.75}[\text{Li}_{0.25}\text{Mn}_{0.75}]\text{O}_2$  and **(b)**  $\text{Na}_{0.6}[\text{Li}_{0.2}\text{Mn}_{0.8}]\text{O}_2$  refined using the  $P6_3/mmc$  space group, which excludes superstructure ordering. Rietveld refinement was performed with

GSAS II software. Refinement parameters are given in the tables in the lower part of the figure, including the goodness of fit (G.O.F.). Inductively coupled plasma (ICP) optical emission spectroscopy was used to confirm the chemical compositions.

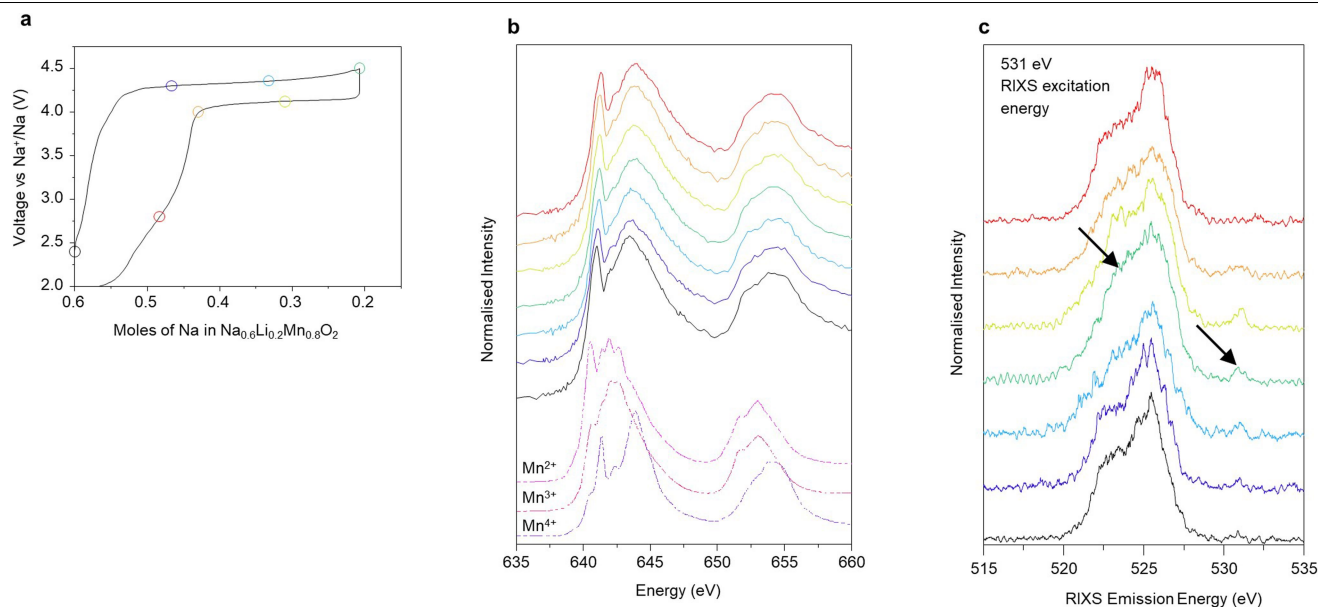


**Extended Data Fig. 2 | Diffraction peaks arising from ribbon superstructure ordering in  $\text{Na}_{0.6}[\text{Li}_{0.2}\text{Mn}_{0.8}]\text{O}_2$ .** **a**, PXRD data for pristine  $\text{Na}_{0.6}[\text{Li}_{0.2}\text{Mn}_{0.8}]\text{O}_2$  indexed using the  $P6_3/mmc$  space group which does not account for superstructure peaks arising from in-plane ordering in the TM layer. **b**, Superstructure region of PXRD compared with computer generated

diffraction patterns. Model crystal structures were prepared with different alignments of ribbon ordered layers. The only structure to successfully match all the peaks is the  $P2_1/c$  space group. Structures are all viewed along the  $[010]$  direction.



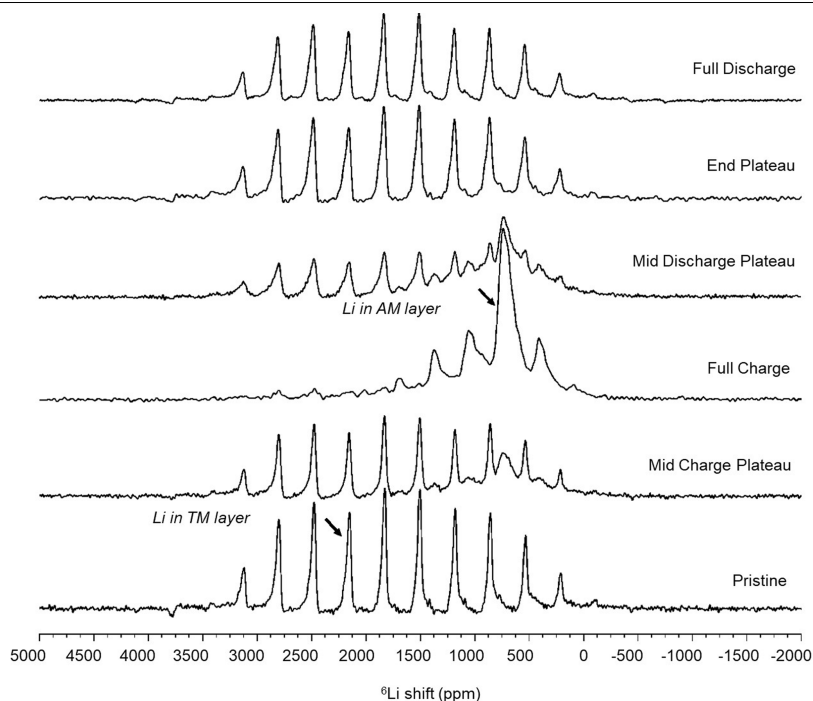
**Extended Data Fig. 3 | Operando gas evolution analysis.** OEMS collected on  $\text{Na}_{0.6}[\text{Li}_{0.2}\text{Mn}_{0.8}]\text{O}_2$  at  $10 \text{ mA g}^{-1}$  between 2 V and 4.5 V. No direct  $\text{O}_2$  loss is observed. Only a very small quantity of  $\text{CO}_2$  is released at 3.5–4.2 V, characteristic of alkali carbonate decomposition; a small amount is released at 4.5 V due to direct electrolyte oxidation. Overall, 0.005 moles of  $\text{CO}_2$  per mole of active material were detected during charge compared with 0.4 moles of charge stored per mole of active material. Even if all of this  $\text{CO}_2$  arose from O loss from the lattice, it would constitute only a minimal contribution (0.02 moles of charge stored per formula unit, f.u., or about 5%), to the charge capacity observed.



**Extended Data Fig. 4 | Manganese L-edge spectra and low-resolution RIXS for  $\text{Na}_{0.6}[\text{Li}_{0.2}\text{Mn}_{0.8}]\text{O}_2$ .** **a**, Electrochemical load curve for first cycle of  $\text{Na}_{0.6}[\text{Li}_{0.2}\text{Mn}_{0.8}]\text{O}_2$  showing state of charge points selected for ex situ analysis. **b**, Manganese L-edge data collected in inverse partial fluorescence yield mode show that  $\text{Na}_{0.6}[\text{Li}_{0.2}\text{Mn}_{0.8}]\text{O}_2$  remains unchanging at  $\text{Mn}^{4+}$  throughout the charge and discharge cycle. Standards shown below are  $\text{MnO}$  (+2),  $\text{Mn}_2\text{O}_3$  (+3) and  $\text{Li}_2\text{MnO}_3$  (+4). **c**, Low-resolution RIXS spectra collected at BL27SU, Spring8

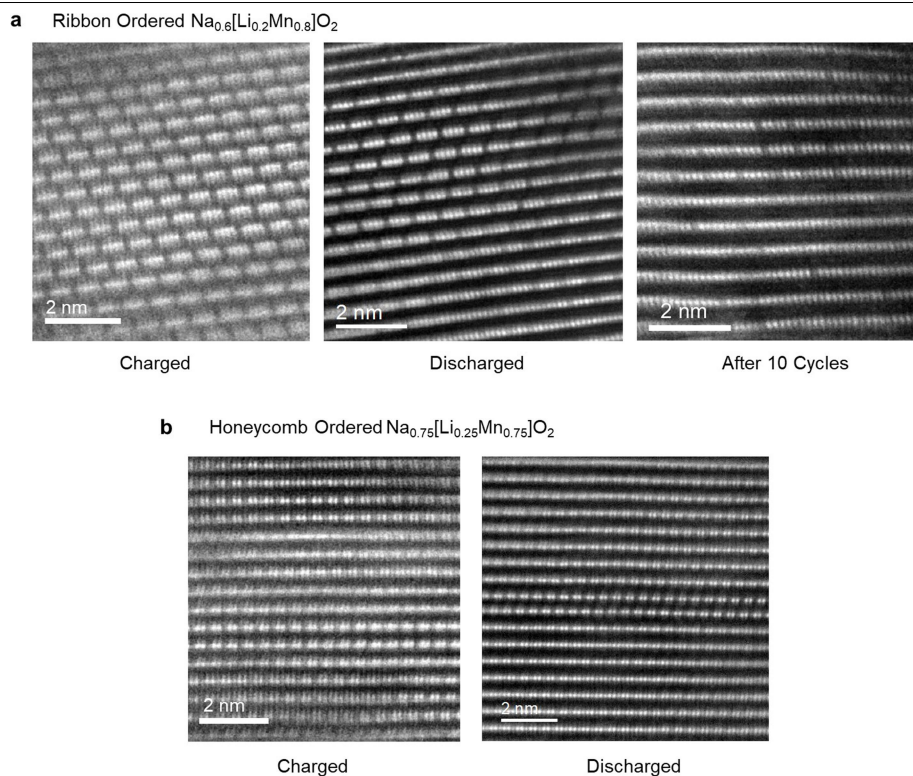
synchrotron, Japan, at 531 eV excitation energy show a new feature at an emission energy of approximately 523 eV corresponding to new hole states formed on O and an increase in the elastic peak intensity (labelled with arrows). These new features disappear on discharge indicating O reduction, and the spectra are almost superimposable with those collected for the pristine material. The intensity of both features appears much less pronounced than the O redox features measured on honeycomb-ordered O-redox materials.





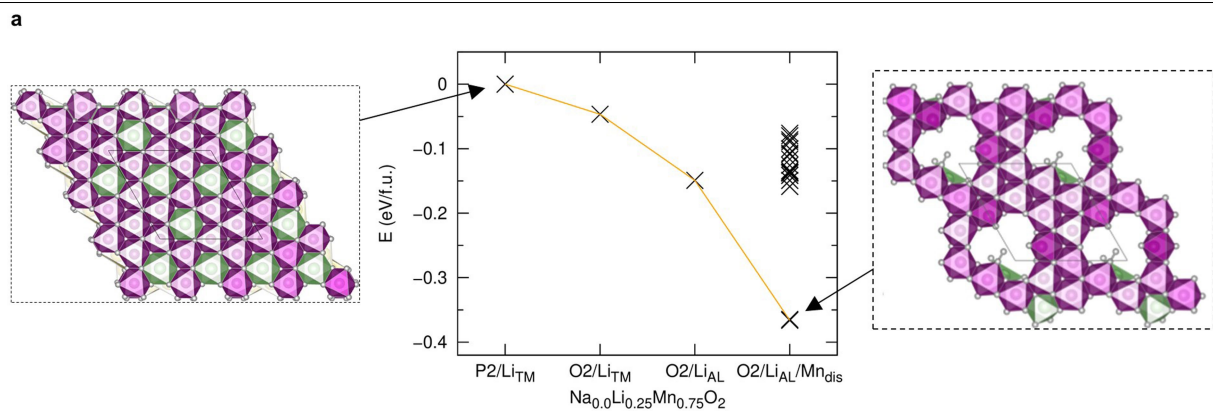
**Extended Data Fig. 5 | Two-phase evolution in Li environment in  $\text{Na}_{0.6}[\text{Li}_{0.2}\text{Mn}_{0.8}]\text{O}_2$ .** Ex situ  $^6\text{Li}$  MAS NMR spectra for  $\text{Na}_{0.6}[\text{Li}_{0.2}\text{Mn}_{0.8}]\text{O}_2$  collected at different states of charge illustrate the two-phase nature of the

charge-discharge plateau. Negligible change in the lithium environment is observed below the end of the plateau on discharge in the single-phase region. Arrows indicate the unique isotropic chemical shifts for Li.



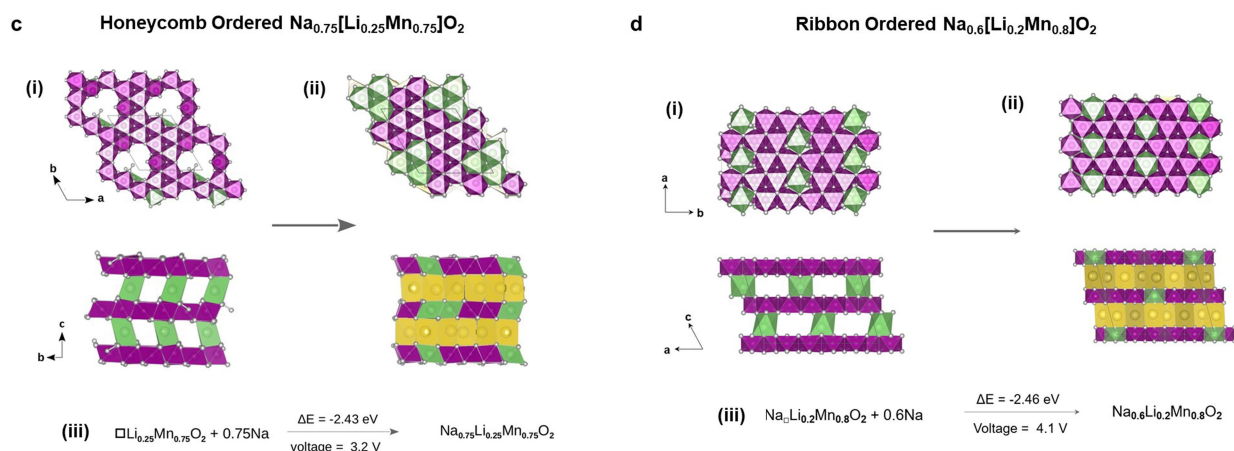
**Extended Data Fig. 6 | Further electron microscopy imaging showing retention of ribbon and loss of honeycomb ordering. a, b, ADF-STEM images showing further spatial investigation of the charged and discharged samples of**

**(a)**  $\text{Na}_{0.6}[\text{Li}_{0.2}\text{Mn}_{0.8}]\text{O}_2$  with ribbon ordering and **(b)**  $\text{Na}_{0.75}[\text{Li}_{0.25}\text{Mn}_{0.75}]\text{O}_2$  with honeycomb ordering. A further image showing  $\text{Na}_{0.6}[\text{Li}_{0.2}\text{Mn}_{0.8}]\text{O}_2$  after 10 charge–discharge cycles is also included.



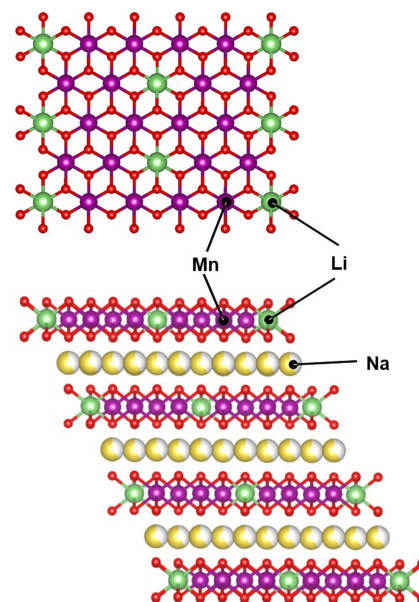
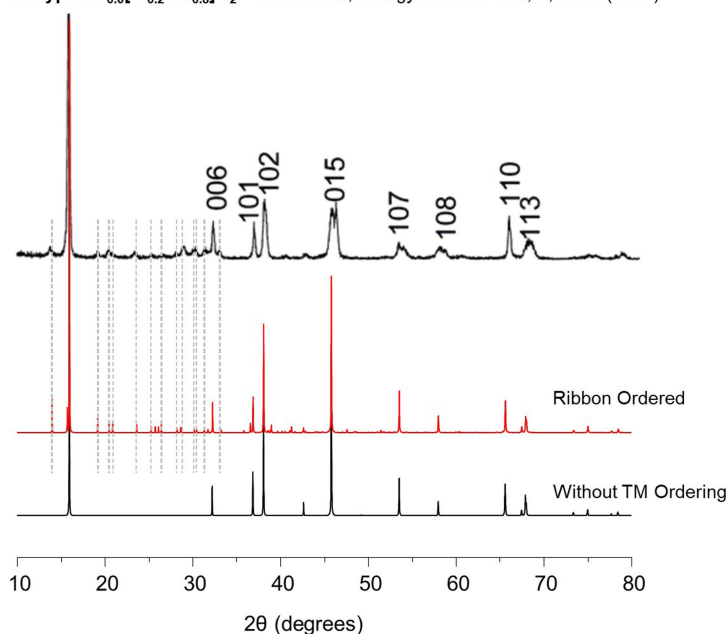
**b**

Model	a (Å)	b (Å)	c (Å)	$\alpha$ (°)	$\beta$ (°)	$\gamma$ (°)
Pristine (DFT)	5.01	5.01	11.04	89.75	90.14	120.16
Pristine (exp.)	4.99	4.99	10.95	90.0	90.0	120.0
Deviation (%)	0.4	0.4	0.8	0.3	0.2	0.1
Charged (DFT)	5.0	5.0	10.38	90.07	89.99	120.1
Charged (exp.)	-	-	10.0	-	-	-
Deviation (%)	-	-	3.8	-	-	-
Discharged (DFT)	5.12	5.15	11.08	89.56	89.82	120.73



**Extended Data Fig. 7 | Energetic stability afforded by  $\text{O}_2$  formation and computed discharge voltage. a.** Energetics of possible configurations of desodiated structural models for  $\text{Na}_{0.0}\text{Li}_{0.25}\text{Mn}_{0.75}\text{O}_2$ . The models considered are: P2-type stacking with Li in the TM layer (P2/ $\text{Li}_{\text{TM}}$ ); O2-type stacking with Li in the TM layer (O2/ $\text{Li}_{\text{TM}}$ ); O2-type stacking with Li in the AL layer (O2/ $\text{Li}_{\text{AL}}$ ); and O2-type stacking with Li in the AL layer and with in-plane Mn disorder (O2/ $\text{Li}_{\text{AL}}/\text{Mn}_{\text{dis}}$ ). In the last case, various Mn disorder configurations were investigated corresponding to the different crosses. The lowest-energy structure is pictured ( $ab$ -plane) and possesses clusters of vacancies and Mn-bound  $\text{O}_2$  with an O–O bond length of 1.2 Å corresponding to molecular  $\text{O}_2$ . For simplicity, the

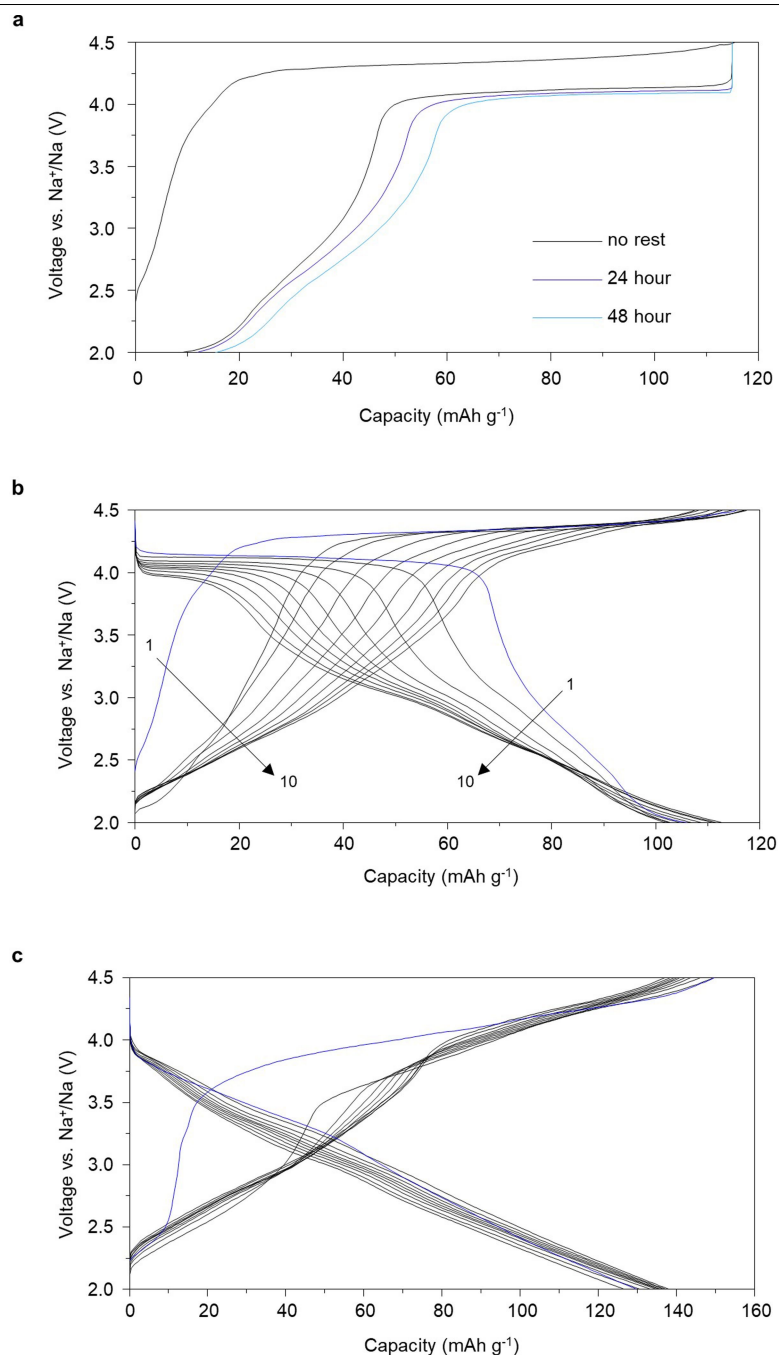
energies of the optimized models are plotted relative to the energy of the model P2/ $\text{Li}_{\text{TM}}$ , the energy of which was set to zero. The yellow curve is a guide to the eye to indicate the models with the lowest total energies. **b.** Calculated lattice parameters for pristine, charged, and discharged  $\text{Na}_{0.75}\text{Li}_{0.25}\text{Mn}_{0.75}\text{O}_2$ . They are compared with experimental data. The deviation between theory and experiment is also reported. **c, d.** Structural models used to compute the change in energy and average voltage for  $\text{Na}_{0.75}\text{Li}_{0.25}\text{Mn}_{0.75}\text{O}_2$  (**c**) and  $\text{Na}_{0.6}\text{Li}_{0.2}\text{Mn}_{0.8}\text{O}_2$  (**d**) respectively. Discharge reactions and calculated voltages are given in (iii). Purple, Mn; green, Li; yellow, Na.



**Extended Data Fig. 8 | Ribbon ordering identified in P3-type  $\text{Na}_{0.6}[\text{Li}_{0.2}\text{Mn}_{0.8}]\text{O}_2$ .** PXRD data for P3-type  $\text{Na}_{0.6}[\text{Li}_{0.2}\text{Mn}_{0.8}]\text{O}_2$  reproduced with permission from ref.<sup>14</sup>. Below are calculated diffraction patterns for the P3 structure with and without ribbon ordering of Li and Mn in the TM layer. Vertical dotted lines show

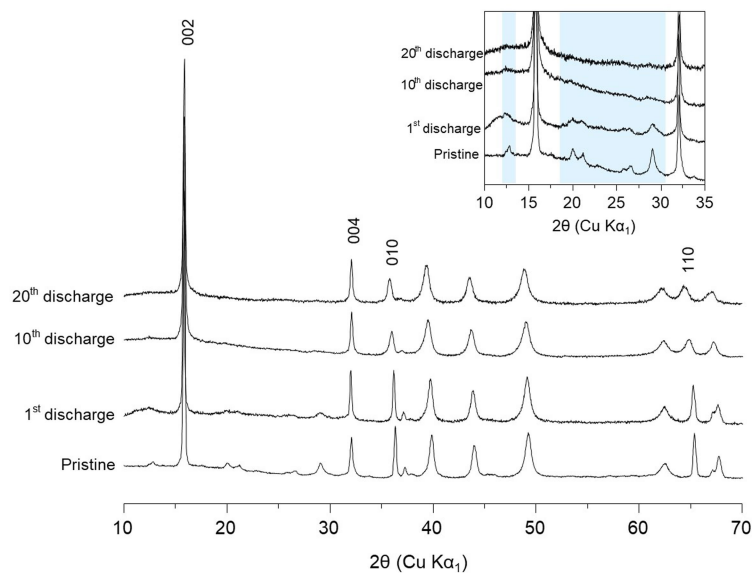
positions of superstructure peaks. The structure used for the calculation of the ribbon-ordered P3-type  $\text{Na}_{0.6}\text{Li}_{0.2}\text{Mn}_{0.6}\text{O}_2$  diffraction pattern is shown to the right and possesses an offset arrangement of ordered layers.





**Extended Data Fig. 9 | Evolution of electrochemical behaviour over resting and cycling.** **a**, Electrochemical load curves for Na<sub>0.6</sub>[Li<sub>0.2</sub>Mn<sub>0.8</sub>]O<sub>2</sub> electrodes charged to 4.5 V at a rate of 10 mA g<sup>-1</sup>, then rested at open circuit voltage (OCV) for varying amounts of time before discharge at 10 mA g<sup>-1</sup>. **b**, **c**, Electrochemical

load curves for ribbon-ordered Na<sub>0.6</sub>[Li<sub>0.2</sub>Mn<sub>0.8</sub>]O<sub>2</sub> and honeycomb-ordered Na<sub>0.75</sub>[Li<sub>0.25</sub>Mn<sub>0.75</sub>]O<sub>2</sub> electrodes respectively, cycled between 2.0 V and 4.5 V at a rate of 10 mA g<sup>-1</sup>. First cycle blue.



**Extended Data Fig. 10 | Gradual loss of ribbon superstructure ordering from diffraction.** Ex situ PXRD patterns for  $\text{Na}_{0.6}[\text{Li}_{0.2}\text{Mn}_{0.8}]\text{O}_2$  in the discharged state (2.0 V) after 1, 10 and 20 charge–discharge cycles (between 2.0 V and 4.5 V at  $10 \text{ mA g}^{-1}$ ). Peaks arising from the periodicity uniquely along the *c*-axis

remain sharp upon cycling (indexed as 002 and 004 according to the  $P6_3/mmc$  space group without superstructure), whereas all other peaks, especially the 010 and 110, which are unique to ordering within the *ab* plane, broaden and reduce in intensity as the ribbon superstructure is lost on cycling.

# Molecular tuning of CO<sub>2</sub>-to-ethylene conversion

<https://doi.org/10.1038/s41586-019-1782-2>

Received: 21 December 2018

Accepted: 1 October 2019

Published online: 20 November 2019

Fengwang Li<sup>1,5</sup>, Arnaud Thevenon<sup>2,5</sup>, Alonso Rosas-Hernández<sup>2,5</sup>, Ziyun Wang<sup>1,5</sup>, Yilin Li<sup>1,5</sup>, Christine M. Gabardo<sup>3</sup>, Adnan Ozden<sup>3</sup>, Cao Thang Dinh<sup>1</sup>, Jun Li<sup>1,3</sup>, Yuhang Wang<sup>1</sup>, Jonathan P. Edwards<sup>3</sup>, Yi Xu<sup>3</sup>, Christopher McCallum<sup>3</sup>, Lizhi Tao<sup>4</sup>, Zhi-Qin Liang<sup>1</sup>, Mingchuan Luo<sup>1</sup>, Xue Wang<sup>1</sup>, Huihui Li<sup>1</sup>, Colin P. O'Brien<sup>3</sup>, Chih-Shan Tan<sup>1</sup>, Dae-Hyun Nam<sup>1</sup>, Rafael Quintero-Bermudez<sup>1</sup>, Tao-Tao Zhuang<sup>1</sup>, Yuguang C. Li<sup>1</sup>, Zhiji Han<sup>2</sup>, R. David Britt<sup>4</sup>, David Sinton<sup>3</sup>, Theodor Agapie<sup>2\*</sup>, Jonas C. Peters<sup>2\*</sup> & Edward H. Sargent<sup>1\*</sup>

The electrocatalytic reduction of carbon dioxide, powered by renewable electricity, to produce valuable fuels and feedstocks provides a sustainable and carbon-neutral approach to the storage of energy produced by intermittent renewable sources<sup>1</sup>. However, the highly selective generation of economically desirable products such as ethylene from the carbon dioxide reduction reaction (CO<sub>2</sub>RR) remains a challenge<sup>2</sup>. Tuning the stabilities of intermediates to favour a desired reaction pathway can improve selectivity<sup>3–5</sup>, and this has recently been explored for the reaction on copper by controlling morphology<sup>6</sup>, grain boundaries<sup>7</sup>, facets<sup>8</sup>, oxidation state<sup>9</sup> and dopants<sup>10</sup>. Unfortunately, the Faradaic efficiency for ethylene is still low in neutral media (60 per cent at a partial current density of 7 milliamperes per square centimetre in the best catalyst reported so far<sup>9</sup>), resulting in a low energy efficiency. Here we present a molecular tuning strategy—the functionalization of the surface of electrocatalysts with organic molecules—that stabilizes intermediates for more selective CO<sub>2</sub>RR to ethylene. Using electrochemical, operando/in situ spectroscopic and computational studies, we investigate the influence of a library of molecules, derived by electro-dimerization of arylpyridiniums<sup>11</sup>, adsorbed on copper. We find that the adhered molecules improve the stabilization of an ‘atop-bound’ CO intermediate (that is, an intermediate bound to a single copper atom), thereby favouring further reduction to ethylene. As a result of this strategy, we report the CO<sub>2</sub>RR to ethylene with a Faradaic efficiency of 72 per cent at a partial current density of 230 milliamperes per square centimetre in a liquid-electrolyte flow cell in a neutral medium. We report stable ethylene electrosynthesis for 190 hours in a system based on a membrane-electrode assembly that provides a full-cell energy efficiency of 20 per cent. We anticipate that this may be generalized to enable molecular strategies to complement heterogeneous catalysts by stabilizing intermediates through local molecular tuning.

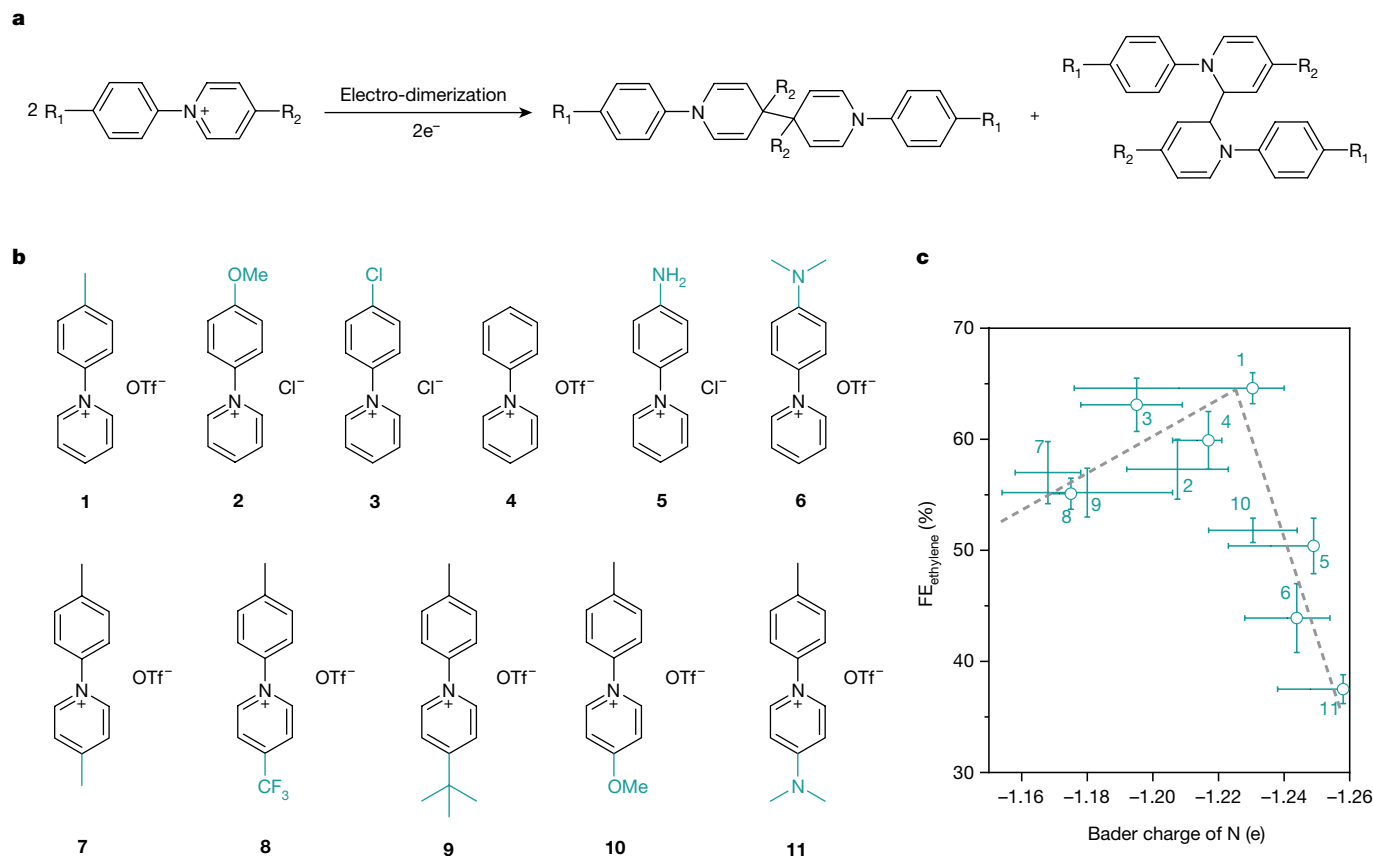
Recently we found that an *N*-aryl-substituted tetrahydro-4,4′-bipyridine organic thin film, formed by reductive electro-dimerization of an *N*-aryl pyridinium additive (Fig. 1a; see Supplementary Information for details), facilitated selective CO<sub>2</sub>RR to multi-carbon products on Cu foils<sup>11</sup>. However, the selectivity and partial current density for ethylene are low (about 40% and 0.5 mA cm<sup>−2</sup>) for practical applications. We sought to clarify factors contributing to the selectivity enhancement to enable further design of new functional molecules with better performance.

Noting that local environment plays a role in electrocatalysis through tuning interactions among reactants/intermediates<sup>12–16</sup>, we postulated

that the *N*-arylpyridinium-derived film may affect the selectivity of CO<sub>2</sub>RR by interacting with the reaction intermediate(s). To test this hypothesis, we first prepared a library of *N*-arylpyridinium salts (**1–11**, Fig. 1b, Supplementary Figs. 1 and 2) expected to display different electronic properties. We then electrodeposited these *N*-arylpyridinium precursors onto a porous polytetrafluoroethylene gas diffusion layer<sup>17</sup> with a sputtered Cu layer serving as both current collector and catalyst. The as-electrodeposited thin film is water-insoluble and consists of a mixture of both constitutional isomers and stereo isomers of *N*-aryl-substituted tetrahydro-bipyridine species (Fig. 1a, Supplementary

<sup>1</sup>Department of Electrical and Computer Engineering, University of Toronto, Toronto, Ontario, Canada. <sup>2</sup>Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA, USA. <sup>3</sup>Department of Mechanical and Industrial Engineering, University of Toronto, Toronto, Ontario, Canada. <sup>4</sup>Department of Chemistry, University of California, Davis, CA, USA.

<sup>5</sup>These authors contributed equally: Fengwang Li, Arnaud Thevenon, Alonso Rosas-Hernández, Ziyun Wang, Yilin Li. \*e-mail: agapie@caltech.edu; jpeters@caltech.edu; ted.sargent@utoronto.ca



**Fig. 1 | Dimerization of *N*-arylpyridinium additives, and correlation of ethylene selectivity with Bader charge.** **a**, Reaction describing the electro-dimerization process that converts an *N*-arylpyridinium salt to a mixture of *N*-aryl-substituted tetrahydro-bipyridines. **b**, Molecular structures of additives **1–11**. OTf<sup>−</sup> is trifluoromethanesulfonate. Cl<sup>−</sup> and OTf<sup>−</sup> are the counter-ions of the derivatives. **c**, Trend for ethylene FE and calculated Bader charge for the nitrogen atom of the *N*-aryl-substituted tetrahydro-bipyridines prepared from **1–11**. Owing to the symmetric molecular structure of the tetrahydro-bipyridines, a hydrogen atom was used to replace half of the dimer unit (see Supplementary Fig. 6 for details). A spread of Bader charges for the nitrogen,

covering the limiting values of the *para,para* and *ortho,ortho* structures, was plotted. The circles correspond to the average contribution from both the *para,para* and *ortho,ortho* isomers where their ratio could be determined by <sup>1</sup>H NMR spectroscopy (see Supplementary Note 1 for details). The error bars for ethylene FE uncertainty represent one standard deviation based on three independent samples. The corresponding error bars for ethylene FE uncertainty were arbitrarily placed in the middle of the limiting values for those tetrahydro-bipyridines for which the *para,para* versus *ortho,ortho* ratio could not be reliably determined by <sup>1</sup>H NMR spectroscopy.

Note 1, Supplementary Figs. 3–5). As expected, Bader charge analysis points to different electron donating abilities of these tetrahydro-bipyridines (Supplementary Fig. 6). Coating of the tetrahydro-bipyridine film onto the Cu electrode does not substantially change its morphology, crystallinity, electronics or wettability, nor does it retard the transport of reactants, ions and products, which is needed in electrocatalytic processes (Supplementary Note 2, Supplementary Figs. 7–10).

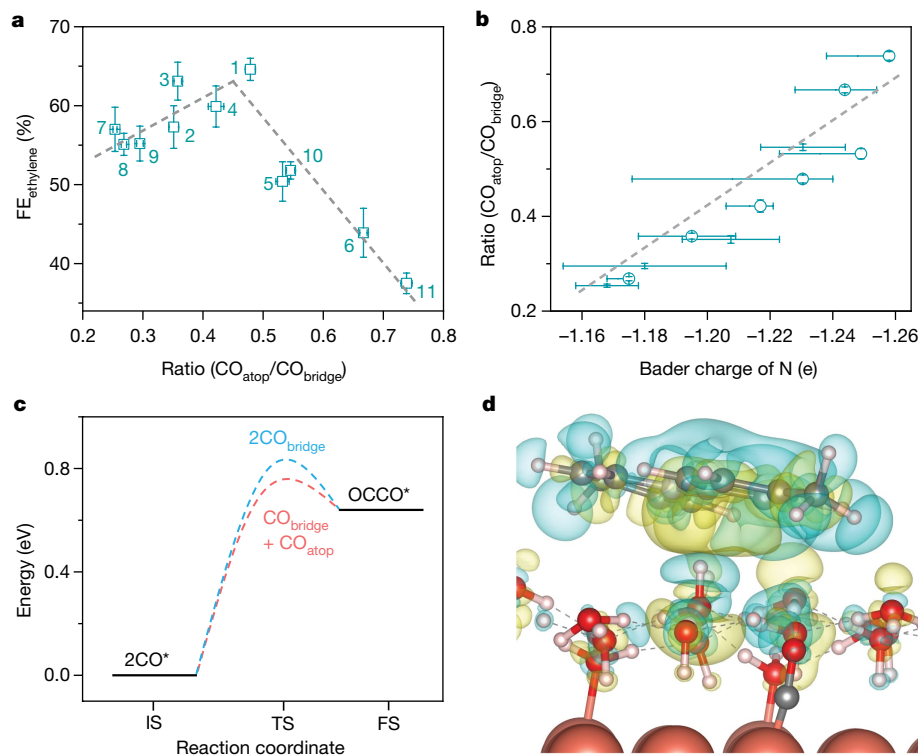
We evaluated CO<sub>2</sub>RR properties of these tetrahydro-bipyridine-functionalized electrodes in a liquid-electrolyte flow cell system (Supplementary Fig. 11), using CO<sub>2</sub>-saturated 1 M aqueous KHCO<sub>3</sub> as the supporting electrolyte. In this system, the abundant catalyst/electrolyte/CO<sub>2</sub> triple-phase interfaces overcome the CO<sub>2</sub> mass-transport limit<sup>17,18</sup> and thus enable commercially relevant current densities<sup>19,20</sup>. We note that, although the large achievable current densities in the flow cell drive up local pH (Supplementary Fig. 12), the tetrahydro-bipyridine layer does not create a further pH gradient near the active Cu surface (Supplementary Note 2). The layer is chemically robust to the locally alkaline environment (Supplementary Fig. 13). The Faradaic efficiency (FE) for ethylene (Supplementary Table 1) on additive-modified Cu-*x* electrodes (*x* = **1–11**), at the optimal applied potentials, −0.82 V to −0.84 V versus the reversible hydrogen electrode (RHE; all potentials are with respect to this reference), was plotted against the Bader charge of the nitrogen atom of each tetrahydro-bipyridine structure (Fig. 1c).

We found a volcano-shaped trend relating FE and Bader charge, with the tetrahydro-bipyridine of moderate electron-donating ability showing the highest ethylene selectivity.

We further found a volcano-shaped relationship between the ethylene selectivity and the ratio of atop-bound CO (CO<sub>atop</sub>) to bridge-bound CO (that is, CO bound to two Cu atoms, hereafter CO<sub>bridge</sub>) on Cu-*x* surfaces (Fig. 2a). We identified and quantified these bound CO configurations through in situ Raman spectroscopic interrogation<sup>21–24</sup> of these surfaces (Supplementary Note 3, Supplementary Figs. 14 and 15, Supplementary Table 2). In all cases, the ratio of CO<sub>atop</sub> to CO<sub>bridge</sub> on Cu-*x* was increased relative to that on bare Cu. Noting a correlation between ethylene selectivity and electron-donation propensity (Fig. 1c), we hypothesized that the change of the relative population of CO<sub>atop</sub> and CO<sub>bridge</sub> could arise from the difference in electron-donating abilities of the tetrahydro-bipyridines. Indeed, we found that the ratio of CO<sub>atop</sub> to CO<sub>bridge</sub> was positively correlated with the Bader charge of the nitrogen atom in the tetrahydro-bipyridines (Fig. 2b). This finding suggests that electron donation to the \*CO stabilizes the atop CO more than it does the bridge CO.

To gain molecular-level insight into the effect of CO binding, we calculated, using density functional theory (DFT), reaction barriers for the CO dimerization step, a critical step along the pathway to C<sub>2</sub> products<sup>5</sup> (that is, products with two carbon atoms, such as ethylene





**Fig. 2 | Mechanistic investigations of the stabilization of CO-bound intermediates.** **a**, The relationship between the ethylene FE and the ratio of atop CO and bridge CO on Cu-**x** electrodes. The relative population of these two kinds of Cu-bound CO was calculated through the integrated areas of each band in the Raman spectra, which are proportional to the corresponding <sup>\*</sup>CO coverage (see Supplementary Note 3 for more details). The error bars for ethylene FE uncertainty represent one standard deviation based on three independent samples. **b**, The relationship between the ratio of atop CO to bridge CO on Cu-**x** and the Bader charge for the nitrogen atom of the *N*-aryl-substituted tetrahydro-bipyridine formed from additive **x**. The Bader charges

and associated uncertainty were calculated using the same protocol as in Fig. 1. The error bars for the ratio of CO<sub>atop</sub> to CO<sub>bridge</sub> in **a** and **b** represent one standard deviation based on two independent measurements. **c**, Energy barriers of the dimerization of two CO at both bridge sites and two CO at bridge and atop sites, respectively. IS, initial state; TS, transient state; FS, final state. **d**, Plots of electron density difference for the CO adsorption with one water layer and the tetrahydro-bipyridine formed from **1**. The yellow and blue contours represent electron density accumulations and depressions, respectively. Dashed lines indicate hydrogen bond network. Red, O; grey, C; blue, N; white, H; pink, Cu.

and ethanol), on Cu(111) with the initial configurations of two <sup>\*</sup>CO on the atop:atop, atop:bridge and bridge:bridge sites (Fig. 2c, Supplementary Fig. 16). We found the lowest barrier of CO dimerization to be at the atop:bridge site with a barrier of 0.72 eV. In comparison, the barrier for the bridge:bridge site is 0.82 eV. The barrier for the atop:atop site could not be identified: one of the CO on atop site tends to relocate to bridge site, suggesting that atop:atop is not favourable for CO dimerization. These findings indicate that neither too large nor too small a population of atop CO favours C<sub>2</sub> selectivity.

We further calculated the adsorption of CO on Cu(111) (Supplementary Fig. 17, Supplementary Table 3). On bare Cu(111), the bridge site appears to be the most stable adsorption site for CO. In the presence of the tetrahydro-bipyridine formed from **1**, the adsorption of CO on both bridge and (especially) atop sites is enhanced, and the atop site becomes favoured compared with the bridge site. The enhancement of CO binding energy decreases the desorption of <sup>\*</sup>CO and increases the likelihood of further reduction of <sup>\*</sup>CO to ethylene (Supplementary Figs. 18–20).

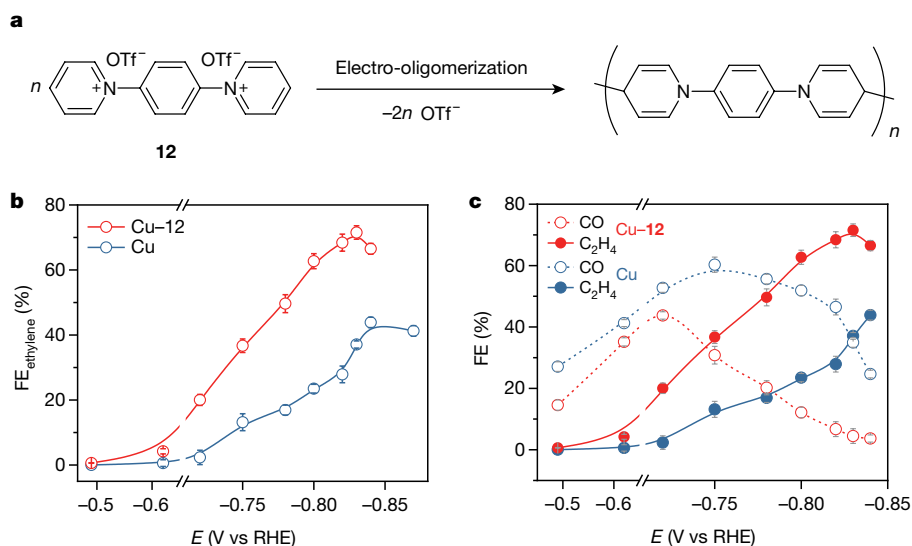
We visualized the interaction between the tetrahydro-bipyridine molecule and <sup>\*</sup>CO through the electron density difference plot (Fig. 2d). The electron density appears to transfer from the molecule to nearby water molecules, changing the electronic distributions of water surrounding <sup>\*</sup>CO, and enhancing CO adsorption in the favourable atop site.

In sum, our working model is that H<sub>2</sub>O-mediated electron density transfer of the tetrahydro-bipyridine film to <sup>\*</sup>CO stabilizes this intermediate, especially on the atop site, and therefore promotes the energy-favourable dimerization of bridge:atop bound CO, leading to enhanced

ethylene selectivity. However, too strong an adsorption of CO caused by strong electron donation of some tetrahydro-bipyridines (right side of the volcano plot in Fig. 1c) results in overload of atop-bound CO and thus yields energy barriers too large for further reaction.

We found, by using operando X-ray absorption spectroscopy (XAS, Supplementary Fig. 21), that tetrahydro-bipyridine does not modulate the oxidation state or coordination environment of Cu—although such modulation is known to promote ethylene formation<sup>9,25</sup>. We also found, from in situ electrochemical electron paramagnetic resonance spectroscopic (EPR) and isotopic labelling studies (Supplementary Figs. 22–24), that tetrahydro-bipyridine does not mediate electron transfers via its conversion to pyridinium radicals<sup>16,26</sup>, nor does it mediate hydrogen-transfer steps.

Because the nitrogen atom of the *N*-aryl-substituted pyridine ring influences the binding of <sup>\*</sup>CO, we posited that an *N*-aryl-pyridinium-derived molecule with more nitrogen sites and optimal electron-donating properties would stabilize more <sup>\*</sup>CO on the Cu surface. Accordingly, we synthesized an *N,N'*-(1,4-phenylene)bispypyridinium salt (**12**, Fig. 3a, Supplementary Fig. 1). In contrast with **1–11**, **12** underwent oligomerization to form an *N*-aryl-dihydropyridine-based oligomer under electrodeposition (Fig. 3a, Supplementary Fig. 5). The Bader charge of the nitrogen atom of the oligomer (Supplementary Fig. 6) is close to that of the tetrahydro-bipyridine from **1**, and, as expected, the ratio of CO<sub>atop</sub> to CO<sub>bridge</sub> on Cu-**12** (Supplementary Fig. 15, Supplementary Table 2) is also close to that on Cu-**1**. Based on the working hypotheses presented here, these findings suggest the Cu-**12** catalyst should approach the top of the volcano plot.



**Fig. 3 | CO<sub>2</sub>RR performance in liquid-electrolyte flow cells. a**, Reaction describing the electro-oligomerization of the *N,N'*-(1,4-phenylene) bispyridinium salt **12** to form an *N*-aryl-dihydropyridine-based oligomer. **b**, FE of ethylene on Cu and Cu-**12** using CO<sub>2</sub>-saturated 1 M KHCO<sub>3</sub> as the supporting

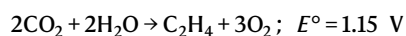
electrolyte. **c**, FEs of CO and ethylene on Cu and Cu-**12** at the applied potential range of -0.47 V to -0.84 V. The error bars for FE uncertainty represent one standard deviation based on three independent samples.

We evaluated the CO<sub>2</sub>RR performance of Cu-**12** in the same flow cell system. The ethylene FE on Cu-**12** is higher than that on bare Cu and other Cu-*x* across the entire applied potential range (-0.49 V to -0.84 V) and achieves a peak value of 72% at -0.83 V (Fig. 3b, Supplementary Tables 1 and 4), higher than previous selectivities reported for ethylene in neutral media (Supplementary Table 5). In contrast, the ethylene FE on bare Cu under similar conditions is below 40%. High selectivity and high current density combine for an ethylene production current of 232 mA cm<sup>-2</sup> at -0.83 V (Supplementary Fig. 25).

We examined the FEs of CO and ethylene across the applied potential range. Although the FE of CO follows the same trend of peaking at moderate potentials, more CO is converted to ethylene on Cu-**12** than on pure Cu (Fig. 3c, Supplementary Table 4). Specifically, at the applied potential of -0.83 V, the FEs of CO and ethylene on Cu-**12** electrode are 5% and 72%, respectively, whereas the values on bare Cu are 35% and 37%, respectively (Supplementary Fig. 25). The FEs of other CO<sub>2</sub>RR products remain similar on both catalysts. These findings suggest that the increased ethylene selectivity arises primarily at the expense of CO evolution. This behaviour agrees with the *in situ* Raman spectroscopy and DFT calculations, where the \*CO is well stabilized for ongoing reduction on the molecularly functionalized Cu electrode.

We confirmed by isotopic CO<sub>2</sub> studies (Supplementary Fig. 26) that the products were from CO<sub>2</sub>RR.

To evaluate the potential of the Cu-**12** catalyst for practical applications, we integrated it into a membrane-electrode-assembly device (Supplementary Note 4, Supplementary Figs. 27–34) for electrosynthesis of ethylene through the overall reaction:

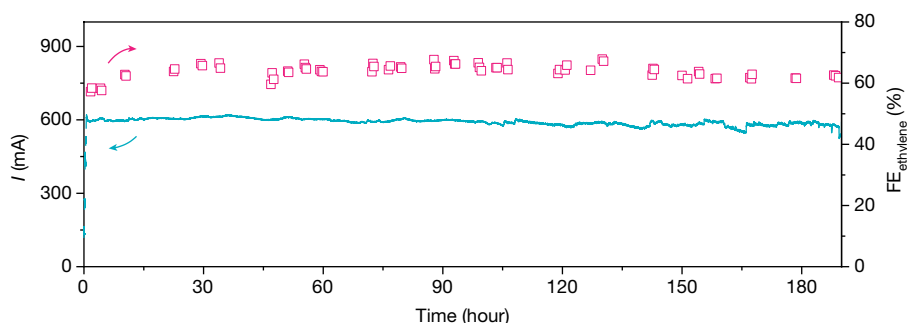


where *E* is the equilibrium potential for the reaction.

We operated the membrane-electrode-assembly system at a full-cell voltage of 3.65 V for 190 h. It exhibited a stable current (approximately 600 mA) and a stable ethylene selectivity (64%) in neutral medium (Fig. 4). The energy efficiency (EE) of the system is determined to be 20% via:

$$\text{EE}_{\text{full-cell}} = (E^\circ \times \text{FE}_{\text{ethylene}}) / E_{\text{full-cell}}$$

Overall, this work presents a strategy to tune the stabilization of intermediates on heterogeneous electrocatalysts through the introduction of organic molecules. Using this strategy, implemented with



**Fig. 4 | Ethylene electrosynthesis in a membrane-electrode assembly device.** The operating current and ethylene FE were monitored for the device. Cu-**12** and iridium oxide supported on titanium mesh were used as the cathode and anode, respectively. Humidified CO<sub>2</sub> was flowed through the gas channels in the cathode, and 0.1 M aqueous KHCO<sub>3</sub> solution was flowed through channels in the anode. The anode and cathode were separated by an anion exchange

membrane to form the membrane-electrode assembly. The total geometric area of the flow field in the cathode is 5 cm<sup>2</sup>, of which 45% is the gas channel while the remaining 55% is the land area (Supplementary Figs. 27 and 28). Full-cell voltage was gradually increased from 3 V to 3.65 V and kept constant starting at time 0.

*N*-aryl-substituted tetrahydro-bipyridine films and a related oligomeric film on a Cu catalyst, we achieved CO<sub>2</sub>-to-ethylene conversion with an ethylene FE of 72% and a full-cell energy efficiency of 20% in neutral media. In light of this performance, in combination with the long-term operating stability, this is a promising strategy for the use of renewable electricity to convert CO<sub>2</sub> into value-added chemicals, thus storing the renewable energy (solar, wind) in the form of chemical energy.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-019-1782-2>.

- Seh, Z. W. et al. Combining theory and experiment in electrocatalysis: insights into materials design. *Science* **355**, eaad4998 (2017).
- De Luna, P. et al. What would it take for renewably powered electrosynthesis to displace petrochemical processes? *Science* **364**, eaav3506 (2019).
- Li, Y. & Sun, Q. Recent advances in breaking scaling relations for effective electrochemical conversion of CO<sub>2</sub>. *Adv. Energy Mater.* **6**, 1600463 (2016).
- Calle-Vallejo, F. & Koper, M. T. Theoretical considerations on the electroreduction of CO to C<sub>2</sub> species on Cu(100) electrodes. *Angew. Chem. Int. Ed.* **52**, 7282–7285 (2013).
- Montoya, J. H., Shi, C., Chan, K. & Nørskov, J. K. Theoretical insights into a CO dimerization mechanism in CO<sub>2</sub> electroreduction. *J. Phys. Chem. Lett.* **6**, 2032–2037 (2015).
- Yang, K. D. et al. Morphology-directed selective production of ethylene or ethane from CO<sub>2</sub> on a Cu mesopore electrode. *Angew. Chem. Int. Ed.* **56**, 796–800 (2017).
- Li, C. W., Ciston, J. & Kanan, M. W. Electroreduction of carbon monoxide to liquid fuel on oxide-derived nanocrystalline copper. *Nature* **508**, 504–507 (2014).
- Jiang, K. et al. Metal ion cycling of Cu foil for selective C–C coupling in electrochemical CO<sub>2</sub> reduction. *Nat. Catal.* **1**, 111–119 (2018).
- Mistry, H. et al. Highly selective plasma-activated copper catalysts for carbon dioxide reduction to ethylene. *Nat. Commun.* **7**, 12123 (2016).
- Zhou, Y. et al. Dopant-induced electron localization drives CO<sub>2</sub> reduction to C<sub>2</sub> hydrocarbons. *Nat. Chem.* **10**, 974–980 (2018).
- Han, Z., Kortlever, R., Chen, H. Y., Peters, J. C. & Agapie, T. CO<sub>2</sub> reduction selective for C<sub>2</sub> products on polycrystalline copper with N-substituted pyridinium additives. *ACS Cent. Sci.* **3**, 853–859 (2017).
- Rosen, B. A. et al. Ionic liquid-mediated selective conversion of CO<sub>2</sub> to CO at low overpotentials. *Science* **334**, 643–644 (2011).
- Masel, R. I. & Rosen, B. A. Catalyst mixtures. US patent 8,956,990 (2015).
- Masel, R. I. & Rosen, B. A. Electrochemical devices comprising novel catalyst mixtures. US patent 9,464,359 (2016).
- Masel, R. I. & Rosen, B. A. Catalyst mixtures. US patent 9,566,574 (2017).
- Barton Cole, E. et al. Using a one-electron shuttle for the multielectron reduction of CO<sub>2</sub> to methanol: kinetic, mechanistic, and structural insights. *J. Am. Chem. Soc.* **132**, 11539–11551 (2010).
- Dinh, C.-T. et al. CO<sub>2</sub> electroreduction to ethylene via hydroxide-mediated copper catalysis at an abrupt interface. *Science* **360**, 783–787 (2018).
- Li, J. et al. Efficient electrocatalytic CO<sub>2</sub> reduction on a three-phase interface. *Nat. Catal.* **1**, 592–600 (2018).
- Ma, S. et al. One-step electrosynthesis of ethylene and ethanol from CO<sub>2</sub> in an alkaline electrolyzer. *J. Power Sources* **301**, 219–228 (2016).
- Jouny, M., Luc, W. W. & Jiao, F. General techno-economic analysis of CO<sub>2</sub> electrolysis systems. *Ind. Eng. Chem. Res.* **57**, 2165–2177 (2018).
- Sheppard, N. & Nguyen, T. T. in *Advances in Infrared and Raman Spectroscopy* Vol. 5 (eds Hawes Clark, R. J. & Hester, R. E.) 67 (Heyden, 1978).
- Gunathunge, C. M. et al. Spectroscopic observation of reversible surface reconstruction of copper electrodes under CO<sub>2</sub> reduction. *J. Phys. Chem. C* **121**, 12337–12344 (2017).
- Heyes, J., Dunwell, M. & Xu, B. CO<sub>2</sub> reduction on Cu at low overpotentials with surface-enhanced in situ spectroscopy. *J. Phys. Chem. C* **120**, 17334–17341 (2016).
- Akemann, W. & Otto, A. Vibrational modes of CO adsorbed on disordered copper films. *J. Raman Spectrosc.* **22**, 797–803 (1991).
- Xiao, H., Goddard, W. A., Cheng, T. & Liu, Y. Cu metal embedded in oxidized matrix catalyst to promote CO<sub>2</sub> activation and CO dimerization for electrochemical reduction of CO<sub>2</sub>. *Proc. Natl Acad. Sci. USA* **114**, 6685–6688 (2017).
- Cole, E. B., Sivasankar, N., Parajuli, R. & Keets, K. A. Reducing carbon dioxide to products. US patent 8,845,878 (2014).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2019

## Data availability

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

**Acknowledgements** This work was financially supported by the Ontario Research Fund: Research Excellence Program, the Natural Sciences and Engineering Research Council (NSERC) of Canada, the CIFAR Bio-Inspired Solar Energy program and the Joint Centre of Artificial Synthesis, a DOE Energy Innovation Hub, supported through the Office of Science of the US Department of Energy under award no. DE-SC0004993. All DFT computations were performed on the IBM BlueGene/Q supercomputer with support from the Southern Ontario Smart Computing Innovation Platform (SOSCIP). SOSCIP is funded by the Federal Economic Development Agency of Southern Ontario, the Province of Ontario, IBM Canada Ltd, Ontario Centres of Excellence, Mitacs and 15 Ontario academic member institutions. This research was enabled in part by support provided by Compute Ontario ([www.computeontario.ca](http://www.computeontario.ca)) and Compute Canada ([www.computecanada.ca](http://www.computecanada.ca)). This research used synchrotron resources of the Advanced Photon Source (APS), an Office of Science User Facility operated for the US Department of Energy (DOE) Office of Science by Argonne National Laboratory, and was supported by the US DOE under contract no. DE-AC02-06CH11357, and the Canadian Light Source and its funding partners. We thank T. Wu and L. Ma for technical support at 9BM beamline of APS. D.S. acknowledges the NSERC E.W.R. Steacie Memorial Fellowship. A.T. acknowledges Marie Skłodowska-Curie Fellowship H2020-MSCA-IF-2017 (793471). J.L. acknowledges the Banting postdoctoral fellowship. C.M.G. acknowledges NSERC for funding in the form of a postdoctoral fellowship from the government of Canada. J.P.E. thanks NSERC, Hatch and the Government of Ontario for their support through graduate scholarships.

**Author contributions** E.H.S., T.A. and J.C.P. supervised this project. F.L. and Y.L. carried out electrochemical experiments. A.T. and A.R.-H. carried out molecule synthesis and characterizations. Z.W. carried out DFT calculations. C.M.G. and F.L. conducted in situ Raman measurement. F.L. and A.O. carried out the membrane-electrode-assembly experiments. J.L. and F.L. performed XAS measurements. Y.W. carried out scanning electron microscopy and electrochemical impedance spectroscopy measurements. J.P.E. measured the contact angle. C.M. carried out the Comsol modelling. L.T. carried out EPR measurement under the supervision of R.D.B. M.L. performed part of the electrochemical experiments. Z.-Q.L., X.W. and H.L. provided help in NMR analysis. C.M.G., C.P.O. and Y.X. provided help in membrane-electrode assembly measurements. C.-S.T. carried out AFM measurement. D.-H.N. conducted X-ray diffraction measurement. R.Q.-B. carried out X-ray photoelectron spectroscopy measurement. C.T.D., T.T.Z., Y.C.L. and Z.H. provided help in materials synthesis and characterizations. F.L. and E.H.S. wrote the manuscript. All authors discussed the results and assisted during manuscript preparation.

**Competing interests** The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41586-019-1782-2>.

**Correspondence and requests for materials** should be addressed to T.A., J.C.P. or E.H.S.

**Peer review information** *Nature* thanks Muhyun Baik, Richard I. Masel and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>.



# Global-scale human impact on delta morphology has led to net land area gain

<https://doi.org/10.1038/s41586-019-1905-9>

Received: 24 January 2019

Accepted: 27 November 2019

Published online: 22 January 2020

J. H. Nienhuis<sup>1,2,3,4\*</sup>, A. D. Ashton<sup>5</sup>, D. A. Edmonds<sup>6</sup>, A. J. F. Hoitink<sup>3</sup>, A. J. Kettner<sup>7</sup>, J. C. Rowland<sup>8</sup> & T. E. Törnqvist<sup>4</sup>

River deltas rank among the most economically and ecologically valuable environments on Earth. Even in the absence of sea-level rise, deltas are increasingly vulnerable to coastal hazards as declining sediment supply and climate change alter their sediment budget, affecting delta morphology and possibly leading to erosion<sup>1–3</sup>. However, the relationship between deltaic sediment budgets, oceanographic forces of waves and tides, and delta morphology has remained poorly quantified. Here we show how the morphology of about 11,000 coastal deltas worldwide, ranging from small bayhead deltas to mega-deltas, has been affected by river damming and deforestation. We introduce a model that shows that present-day delta morphology varies across a continuum between wave (about 80 per cent), tide (around 10 per cent) and river (about 10 per cent) dominance, but that most large deltas are tide- and river-dominated. Over the past 30 years, despite sea-level rise, deltas globally have experienced a net land gain of  $54 \pm 12$  square kilometres per year (2 standard deviations), with the largest 1 per cent of deltas being responsible for 30 per cent of all net land area gains. Humans are a considerable driver of these net land gains—25 per cent of delta growth can be attributed to deforestation-induced increases in fluvial sediment supply. Yet for nearly 1,000 deltas, river damming<sup>4</sup> has resulted in a severe (more than 50 per cent) reduction in anthropogenic sediment flux, forcing a collective loss of  $12 \pm 3.5$  square kilometres per year (2 standard deviations) of deltaic land. Not all deltas lose land in response to river damming: deltas transitioning towards tide dominance are currently gaining land, probably through channel infilling. With expected accelerated sea-level rise<sup>5</sup>, however, recent land gains are unlikely to be sustained throughout the twenty-first century. Understanding the redistribution of sediments by waves and tides will be critical for successfully predicting human-driven change to deltas, both locally and globally.

River damming and land-use change affect the sediment supply to deltas, and can lead to substantial physical transformations of the coastal landscape. Existing attempts to predict delta morphology are conceptually rich but often qualitative<sup>6–11</sup>. Most prominently, Galloway<sup>7</sup> introduced a process-based ternary diagram, hypothesizing that delta morphology reflects the relative importance of wave, tide and river forcing. However, the lack of a quantitative prediction of delta morphology for a given relative influence of each forcing has prevented direct application of this foundational ternary diagram to understanding delta form. For example, how does decreased sediment supply affect deltas and how can this translate into land gain or land loss? A fundamental limitation in predicting delta change has been the poor understanding of how sediment supply has shaped modern delta

morphology itself, motivating our development of an a priori theory of the controls of delta morphology.

## A new model for delta change

On the basis of two recent quantitative studies<sup>12,13</sup>, we here introduce a ternary diagram that allows prognosis of delta morphology and morphologic change using sediment fluxes (Fig. 1a). We apply this approach on a global scale. First, we predict delta morphology for conditions that resemble a world without substantial human impact on the fluvial sediment supply. Next, we compare these predictions to the delta morphology that is expected on the basis of recent modifications to sediment fluxes due to both deforestation and river damming.

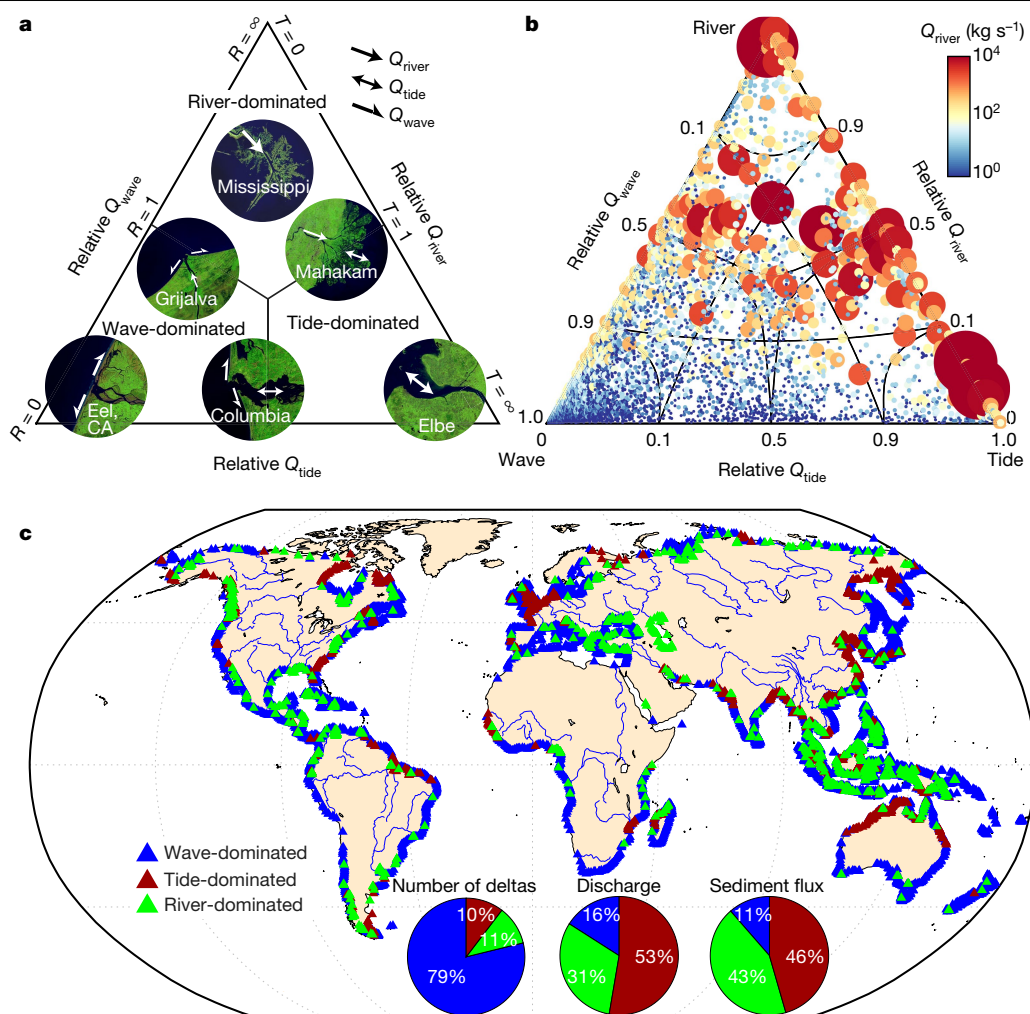
<sup>1</sup>Department of Physical Geography, Utrecht University, Utrecht, The Netherlands. <sup>2</sup>Department of Earth, Ocean, and Atmospheric Sciences, Florida State University, Tallahassee, FL, USA.

<sup>3</sup>Environmental Sciences, Wageningen University and Research, Wageningen, The Netherlands. <sup>4</sup>Department of Earth and Environmental Sciences, Tulane University, New Orleans, LA, USA.

<sup>5</sup>Department of Geology and Geophysics, Woods Hole Oceanographic Institution, Woods Hole, MA, USA. <sup>6</sup>Department of Earth and Atmospheric Sciences, Indiana University, Bloomington, IN, USA.

<sup>7</sup>Institute of Arctic and Alpine Research, University of Colorado, Boulder, CO, USA. <sup>8</sup>Earth & Environmental Sciences Division, Los Alamos National Laboratory, Los Alamos, NM, USA.

\*e-mail: j.h.nienhuis@uu.nl



**Fig. 1 | Global distribution of predicted pristine delta morphologies.**

**a**, Galloway<sup>7</sup> ternary diagram, recast to show the relative sediment fluxes  $Q_{\text{wave}}$ ,  $Q_{\text{tide}}$  and  $Q_{\text{river}}$  (see Methods). Insets are satellite images of representative delta morphologies, with arrows highlighting the predicted direction and magnitude of sediment fluxes. Map imagery in Figs. 1, 3 and Extended Data Fig. 5 from NASA, Google Earth, TerraMetrics, 2019. **b**, Prediction of pristine ( $Q_{\text{river}}^p$ ) morphology of 10,848 deltas sized and coloured by fluvial sediment

flux. Axes follow a sigmoidal, rather than linear, function to better illustrate the distribution of strongly wave-, river- or tide-dominated deltas. **c**, Global geographic distribution of predicted pristine delta morphologies (see .kml file at <https://doi.org/10.17605/OSF.IO/S28QB>). Plots in Figs. 1–3 and Extended Data Figs. 1–5 generated by Matlab 2018b (<https://mathworks.com/products/matlab.html>).

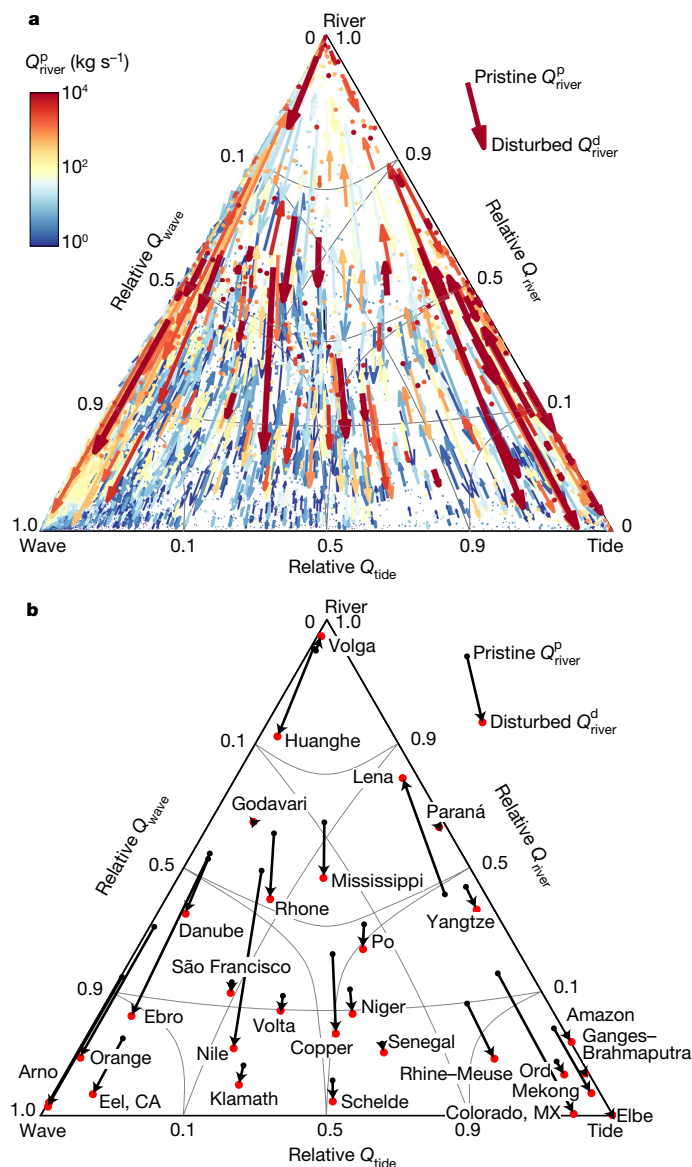
We distinguish between two formative values of the fluvial sediment supply ( $Q_{\text{river}}$ , in kilograms per second), representing pristine sediment fluxes before substantial anthropogenic influences ( $Q_{\text{river}}^p$ ) and contemporary ('disturbed') sediment fluxes accounting for dam construction and land-use change in the contributing drainage basins ( $Q_{\text{river}}^d$ )<sup>14</sup>. Because deltas respond to sediment flux changes over timespans of decades to centuries<sup>15</sup>, our delta morphology predictions based on  $Q_{\text{river}}^d$  correspond to a future equilibrium state towards which deltas are currently evolving. Using observations of delta land area changes in 1985–2015, we can investigate how much humans have changed deltas and how deltas may change in the future.

Our ternary diagram compares the fluvial sediment supply to tide- and wave-driven sediment fluxes near the river mouth. First, in the absence of tides, a delta is expected to attain a wave-dominated, triangular shape if the potential for waves to move sediment away from the river mouth ( $Q_{\text{wave}}$ , in kilograms per second; see Methods) exceeds the delivered fluvial sediment flux ( $Q_{\text{river}}$ ). Importantly,  $Q_{\text{river}}$  and  $Q_{\text{wave}}$  enable predictions independent of the observed delta morphology and allow these sediment fluxes to be used for delta change forecasting. The ratio  $Q_{\text{river}}/Q_{\text{wave}}$  (termed the fluvial dominance ratio,  $R$ ) indicates whether a delta does not deflect the coastline ( $R \approx 0$ ; for example, Eel; Fig. 1a),

has a roughly triangular shape with a shoreline angle between  $0^\circ$  and  $45^\circ$  ( $0 < R < 1$ ; for example, Grijalva), or is river-dominated ( $R > 1$ ; for example, Mississippi). Increases in  $R$  lead to increased deposition near the river mouth, whereas decreases in  $R$  can result in distal shoreline progradation even as the river mouth erodes<sup>12</sup>.

In the absence of waves, delta morphology is determined by the competition between river discharge and tidal flows. Morphologically, tidal dominance manifests itself as a seaward widening of the channel banks<sup>13,16</sup>. By contrast, river-dominated delta channels have an approximately constant width. The tidal dominance ratio  $T$ , as originally defined<sup>13</sup>, relates the tidal discharge amplitude to the mean fluvial discharge. Here we use  $T$  as a ratio of sediment fluxes and define a tidal sediment flux ( $Q_{\text{tide}}$ , in kilograms per second) along with a fluvial sediment flux ( $Q_{\text{river}}$ , in kilograms per second) (Fig. 1a, Methods). If  $T < 1$ , the delta is river-dominated and there is no flow reversal in the deltaic channel(s). If  $T > 1$ , the delta is tide-dominated and the widened deltaic channel(s), or some portion thereof, experience(s) flow reversal. Changes in  $T$  will affect delta channels; for example, a decrease in fluvial sediment flux ( $Q_{\text{river}}$ ) will cause the channel to infill and narrow<sup>13</sup>.

Our ternary diagram represents the relative contribution of  $Q_{\text{tide}}$ ,  $Q_{\text{river}}$  and  $Q_{\text{wave}}$ , and therefore also two morphological attributes of a

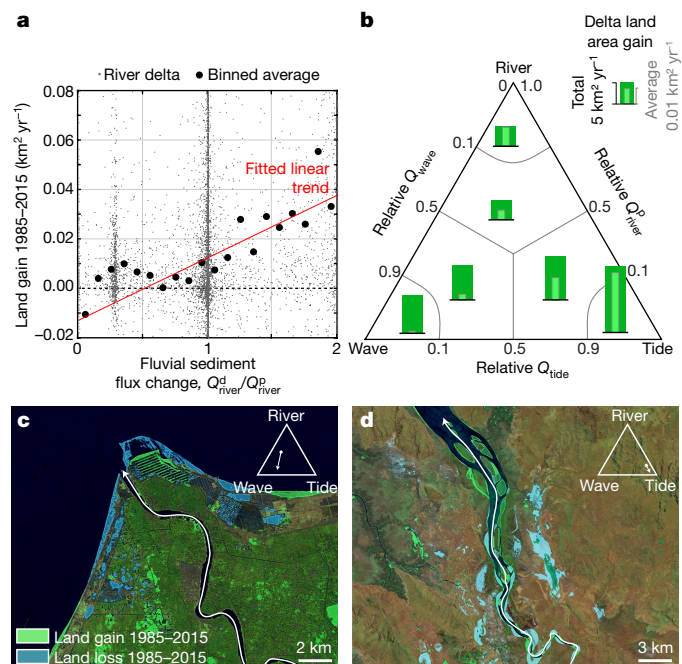


**Fig. 2 | Predicted delta morphologic change from pristine to future equilibrium conditions.** **a**, Arrows indicate the direction and magnitude of the predicted change. Colour and thickness indicate the pristine fluvial sediment flux. **b**, Predicted anthropogenically driven morphologic change for a selection of well-known deltas. See also Extended Data Table 4.

delta: the seaward divergence of the channel banks and the shoreline protrusion angle (Fig. 1a). It allows us to explore delta morphologies that arise from varying  $Q_{\text{tide}}$ ,  $Q_{\text{river}}$  and  $Q_{\text{wave}}$ , including the expected morphology of deltas near the limit of low fluvial sediment flux, now or in the future<sup>17</sup>. Deltas near this limit are often referred to as strand plains (for example, São Francisco<sup>18</sup>) or alluvial estuaries (for example, Elbe<sup>8</sup>). Here we show that this wide variety of coastal morphologies with different sizes lies along a continuum that can be characterized by the relative balance of these three sediment fluxes. For simplicity, we therefore refer to all morphologies within our ternary diagram as deltas—a broader definition compared to other studies<sup>19</sup>.

### A global assessment of delta change

To predict pristine delta morphology globally, we determined the location of coastal deltas worldwide ( $n = 10,848 \pm 494$ ; 2 s.d.) and calculated pristine river-, wave- and tide-driven sediment fluxes. These fluxes



**Fig. 3 | Rates and drivers of delta land area change over the period 1985–2015.** **a**, **b**, Land area change rates related to changes in the fluvial sediment supply (**a**) and pristine delta morphology (**b**). **c**, **d**, Land change in the Nile Delta, Egypt (**c**) and the Ord River Delta, Australia (**d**). Map imagery, NASA, Google Earth, TerraMetrics, 2019 and ref.<sup>20</sup>. The inset diagrams indicate the predicted morphologic change.

occur in all combinations, and the predicted delta morphologies vary across a continuum between wave, tide and river dominance, as tested against observed morphologies (see Methods). Most deltas are wave-dominated ( $\sim 79\% \pm 9\%$ ; 2 s.d.); however, large deltas ( $Q_{\text{river}}^p > 50 \text{ kg s}^{-1}$ ,  $n = 701$ ) are predominantly (68%) river- or tide-dominated (Fig. 1b), owing to their large fluvial sediment flux and their low-gradient delta plains ( $5 \times 10^{-4}$  versus  $3 \times 10^{-3}$  for all deltas on average), making them conducive to large tidal sediment fluxes<sup>13</sup>. River- and tide-dominated deltas are associated with 83% of the modern fluvial discharge and 87% of the modern sediment flux to the global ocean.

A comparison of equilibrium predictions for pristine and disturbed sediment fluxes shows the extent to which humans are likely to be modifying delta morphology by influencing river discharge and sediment fluxes. In total, 970 deltas have had their fluvial sediment supply reduced by  $>50\%$ , collectively from  $-9 \times 10^4 \text{ kg s}^{-1}$  to  $-2 \times 10^4 \text{ kg s}^{-1}$ , resulting in a shift towards wave or tide dominance (Fig. 2a). On the other hand, human-driven soil erosion, mostly through deforestation, is predicted to have caused a  $>50\%$  increase in sediment flux, or  $\sim 5 \times 10^4 \text{ kg s}^{-1}$ , to  $\sim 1,500$  deltas. We predict that sediment supply changes are forcing considerable ongoing adjustments in the shoreline protrusion and channel width of many well-known deltas (Fig. 2b).

Next, we use the Aqua Monitor<sup>20</sup> to investigate how our predicted ongoing morphologic change is reflected in recent delta surface area change (see Methods). We find that over the past 30 years, deltas globally have gained  $181 \pm 8.3 \text{ km}^2 \text{ yr}^{-1}$  and lost  $127 \pm 8.3 \text{ km}^2 \text{ yr}^{-1}$ , resulting in a net gain of  $54 \pm 11.8 \text{ km}^2 \text{ yr}^{-1}$  (2 s.d.). With a combined  $-9 \times 10^9 \text{ m}^3 \text{ yr}^{-1}$  fluvial sediment flux to the global ocean<sup>21</sup>, deltas on average require  $150 \text{ m}^3$  of sediment delivered to the coast for every square metre of land gain. Delta growth is particularly pronounced for tide-dominated deltas, representing 46% of the net land gain.

We find that humans have measurably altered delta growth rates globally (Fig. 3a, Table 1). Human-induced changes to the fluvial sediment flux ( $Q_{\text{river}}^d - Q_{\text{river}}^p$ ) explain 16% of the recent delta land area

**Table 1 | Global delta morphology and morphodynamic change**

	Number of deltas	Total $Q_{\text{river}}^p$ (kg s <sup>-1</sup> )	Total $Q_{\text{river}}^d$ (kg s <sup>-1</sup> )	Land gain (km <sup>2</sup> yr <sup>-1</sup> )	Land loss (km <sup>2</sup> yr <sup>-1</sup> )	Net land gain (km <sup>2</sup> yr <sup>-1</sup> )
Wave-dominated	8,552	$6.0 \times 10^4$	$5.9 \times 10^4$	$35 \pm 7$	$-17 \pm 7$	$19 \pm 10$
River-dominated	1,169	$20 \times 10^4$	$15 \times 10^4$	$49 \pm 3$	$-39 \pm 3$	$10 \pm 4$
Tide-dominated	1,127	$22 \times 10^4$	$22 \times 10^4$	$97 \pm 3$	$-72 \pm 3$	$25 \pm 4$
Fluvial flux decrease (>50%)	970	$9.2 \times 10^4$	$1.8 \times 10^4$	$15 \pm 3$	$-27 \pm 3$	$-12 \pm 4$
Fluvial flux increase (>50%)	1,478	$3.1 \times 10^4$	$7.6 \times 10^4$	$36 \pm 3$	$-11 \pm 3$	$25 \pm 4$
Tidal reworking <sup>a</sup>	234	$4.2 \times 10^4$	$1.0 \times 10^4$	$2 \pm 1$	$-1 \pm 1$	$0.9 \pm 1$
Wave reworking <sup>b</sup>	736	$5.0 \times 10^4$	$0.8 \times 10^4$	$12 \pm 2$	$-25 \pm 2$	$-13 \pm 3$
Largest 1%	108	$35 \times 10^4$	$29 \times 10^4$	$103 \pm 1$	$-88 \pm 1$	$15 \pm 1$
Largest 10%	1,085	$46 \times 10^4$	$40 \times 10^4$	$143 \pm 3$	$-109 \pm 3$	$34 \pm 4$
Largest 100% (all deltas)	10,848	$49 \times 10^4$	$43 \times 10^4$	$181 \pm 8$	$-127 \pm 8$	$54 \pm 12$

Error limits indicate 2 s.d.

<sup>a</sup>Tidal reworking defined as a fluvial sediment flux decrease greater than 50% and  $Q_{\text{wave}} < Q_{\text{tide}}$ .

<sup>b</sup>Wave reworking defined as a fluvial sediment flux decrease greater than 50% and  $Q_{\text{wave}} > Q_{\text{tide}}$ .

changes ( $P = 0$ ). Deforestation has led to land gain, thus far exceeding land loss due to river dams. Delta change is most pronounced in South, Southeast and East Asia, where 57% of all new deltaic land is gained and 61% of all delta land loss occurs. North America, owing to the rapid decline of the Mississippi Delta, partly due to damming<sup>22</sup>, is the only continent with a net decrease in deltaic area (Extended Data Table 3).

Delta response to river damming depends on how waves and tides redistribute (rework) deltaic sediment (Fig. 3b). Two dominant patterns emerge. Deltas that are predicted to become more wave-dominated are, on average, eroding (Table 1). Morphologically, this change is expected because wave reworking of the delta near the river mouth results in erosion<sup>23</sup> (Fig. 3c). However, tidally influenced deltas that face markedly reduced fluvial sediment supply are slightly gaining (or not necessarily losing) land area (Table 1, Fig. 3d). This counterintuitive result is caused by the infilling of deltaic channels<sup>13</sup>. In contrast to some studies (for example, in the Amazon<sup>24</sup> or Yangtze<sup>25</sup>) that assume that dams will lead to delta erosion, our analysis suggests that tides can overcompensate for the reduced fluvial discharge or sediment input and increase landward sediment transport. Increased landward transport probably results from the relative enhancement of tidal flood flow in cases where fluvial discharge (peaks) are decreasing<sup>26,27</sup> and comes at the expense of the extensive subaqueous delta.

## Discussion

Because our predictions of delta morphologic change are global in scale, they exclude various processes affecting deltas now and in the future, such as relative sea-level change and direct anthropogenic modification—processes included in measurements of land area change. For heavily modified delta plains (for example, the Rhine–Meuse Delta), morphologic predictions based on changes in the fluvial sediment flux can indicate long-term system tendencies; however, the actual response will most probably involve direct human–delta interactions not considered by our approach.

Our ternary diagram simplifies delta morphology into two shape metrics: delta protrusion angle and channel width. It therefore differs from earlier, qualitative work. For example, the São Francisco river is often thought of as having an end-member wave-dominated delta<sup>7</sup>. Here we show that the delta is wave-dominated, but that fluvial sediment has created a substantial shoreline protrusion ( $R \approx 0.3$ ) and that tides probably create flow reversal at the river mouth ( $T \approx 1$ ). We note also that two deltas that are placed near each other in our framework (for example, Volga and Huanghe; Fig. 2b) might be considered to be different on the basis of other aspects of delta morphology (for example, shoreline rugosity, number of distributary channels). Our ternary diagram can help explore the origin of such morphologic differences.

For example,  $Q_{\text{river}}$  is split across distributary channels, whereas  $Q_{\text{wave}}$  and  $Q_{\text{tide}}$  act on each river mouth. Via channel bifurcation, deltas that are marginally river-dominated can therefore transition towards wave or tide dominance<sup>12</sup>. Conversely, because  $Q_{\text{wave}}$  suppresses channel bifurcation<sup>28</sup>, we could potentially predict the number of distributary channels for river deltas.

Changes to sediment fluxes explain dominant trends in delta plan-form evolution and are sufficiently general to allow for coupling with other processes. Sea-level rise and subsidence, for example, tend to increase deltaic channel and topset aggradation<sup>29</sup>, which would reduce fluvial sediment supply to the river mouth ( $Q_{\text{river}}$ ) and result in a relative increase of wave and tide dominance. Other controls on delta morphology, such as grain size or wave climate changes<sup>30</sup>, can be incorporated into our model, but appropriate data for global applications are currently lacking. For example, grain size is inversely correlated to  $Q_{\text{tide}}$  and  $Q_{\text{wave}}$  (refs. <sup>12,13</sup>), making coarser-grained deltas more likely to be river-dominated.

In conclusion, we can successfully predict large-scale delta morphology and we find that human intervention in drainage basins has had a considerable global effect. The recent reductions in sediment supply explain important patterns of land loss in cases where waves take over. Yet on a global scale, land gains resulting from deforestation exceed losses due to river damming. In the future, however, dam emplacement and sand mining is projected to accelerate in developing nations, further lowering fluvial sediment supply to river deltas<sup>31,32</sup>. Sea-level rise and land subsidence rates are increasing in many deltas<sup>3,33,34</sup>. Future predictions of delta morphology therefore will need to consider further diminished sediment loads and higher relative sea-level rise rates.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-019-1905-9>.

1. Syvitski, J. P. M. et al. Sinking deltas due to human activities. *Nat. Geosci.* **2**, 681–686 (2009).
2. Tessler, Z. D. et al. Profiling risk and sustainability in coastal deltas of the world. *Science* **349**, 638–643 (2015).
3. Pelletier, J. D. et al. Forecasting the response of Earth's surface to future climatic and land use changes: A review of methods and research needs. *Earth's Future* **3**, 220–251 (2015).
4. Syvitski, J. P. M. & Kettner, A. Sediment flux and the Anthropocene. *Philos. Trans. R. Soc. A* **369**, 957–975 (2011).
5. Nerem, R. S. et al. Climate-change-driven accelerated sea-level rise detected in the altimeter era. *Proc. Natl Acad. Sci. USA* **115**, 2022–2025 (2018).



6. Wright, L. D. & Coleman, J. M. Variations in morphology of major river deltas as functions on ocean wave and river discharge regimes. *Am. Assoc. Pet. Geol. Bull.* **57**, 370–398 (1973).
7. Galloway, W. D. in *Deltas, Models for Exploration* (ed. Broussard, M. L.) 86–98 (Houston Geological Society, 1975).
8. Boyd, R., Dalrymple, R. & Zaitlin, B. A. Classification of clastic coastal depositional environments. *Sedim. Geol.* **80**, 139–150 (1992).
9. Orton, G. J. & Reading, H. G. Variability of deltaic processes in terms of sediment supply, with particular emphasis on grain size. *Sedimentology* **40**, 475–512 (1993).
10. Bhattacharya, J. P. in *Facies Models Revisited* (eds Posamentier, H. W. & Walker, R. G.) 237–292 (Society for Sedimentary Geology, 2006).
11. Leeder, M. R. *Sedimentology and Sedimentary Basins: From Turbulence to Tectonics* (Wiley-Blackwell, 2011).
12. Nienhuis, J. H., Ashton, A. D. & Giosan, L. What makes a delta wave-dominated? *Geology* **43**, 511–514 (2015).
13. Nienhuis, J. H., Houtink, A. J. F. & Törnqvist, T. E. Future change to tide-influenced deltas. *Geophys. Res. Lett.* **45**, 3499–3507 (2018).
14. Cohen, S., Kettner, A. J. & Syvitski, J. P. M. Global suspended sediment and water discharge dynamics between 1960 and 2010: continental trends and intra-basin sensitivity. *Global Planet. Change* **115**, 44–58 (2014).
15. Swenson, J. B. Relative importance of fluvial input and wave energy in controlling the timescale for distributary-channel avulsion. *Geophys. Res. Lett.* **32**, L23404 (2005).
16. Langbein, W. B. The hydraulic geometry of a shallow estuary. *Int. Assoc. Sci. Hydrol. Bull.* **8**, 84–94 (1963).
17. Dalrymple, R. W. in *Facies Models: Response to Sea Level Change* (eds Walker, R. G. & James, N. P.) 195–218 (Geological Association of Canada, 1992).
18. Dominguez, J. M. L. The São Francisco strandplain: a paradigm for wave-dominated deltas? *Geol. Soc. Lond. Spec. Publ.* **117**, 217–231 (1996).
19. Caldwell, R. L. et al. A global delta dataset and the environmental variables that predict delta formation on marine coastlines. *Earth Surf. Dyn.* **7**, 773–787 (2019).
20. Donchyts, G. et al. Earth's surface water change over the past 30 years. *Nat. Clim. Change* **6**, 810–813 (2016).
21. Syvitski, J. P. M., Vörösmarty, C. J., Kettner, A. J., & Green, P. Impact of humans on the flux of terrestrial sediment to the global coastal ocean. *Science* **308**, 376–380 (2005).
22. Couvillion, B. R., Beck, H., Schoolmaster, D. & Fischer, M. *Land Area Change in Coastal Louisiana (1932 to 2016)*. Map 3381 (U.S. Geological Survey, 2017); <https://doi.org/10.3133/sim3381>.
23. Nienhuis, J. H., Ashton, A. D., Roos, P. C., Hulscher, S. J. M. H. & Giosan, L. Wave reworking of abandoned deltas. *Geophys. Res. Lett.* **40**, 5899–5903 (2013).
24. Latrubesse, E. M. et al. Damming the rivers of the Amazon basin. *Nature* **546**, 363–369 (2017).
25. Yang, Z. et al. Dam impacts on the Changjiang (Yangtze) River sediment discharge to the sea: the past 55 years and after the Three Gorges Dam. *Water Resour. Res.* **42**, 04407 (2006).
26. Houtink, A. J. F., Wang, Z. B., Vermeulen, B., Huismans, Y. & Kästner, K. Tidal controls on river delta morphology. *Nat. Geosci.* **10**, 637–645 (2017).
27. Dai, Z., Fagherazzi, S., Mei, X., Chen, J. & Meng, Y. Linking the infilling of the North Branch in the Changjiang (Yangtze) estuary to anthropogenic activities from 1958 to 2013. *Mar. Geol.* **379**, 1–12 (2016).
28. Jerolmack, D. J. & Swenson, J. B. Scaling relationships and evolution of distributary networks on wave-influenced deltas. *Geophys. Res. Lett.* **34**, L23402 (2007).
29. Muto, T. & Swenson, J. B. Large-scale fluvial grade as a nonequilibrium state in linked depositional systems: theory and experiment. *J. Geophys. Res.* **110**, F03002 (2005).
30. Casas-Prat, M., Wang, X. L. & Swart, N. CMIP5-based global wave climate projections including the entire Arctic Ocean. *Ocean Model.* **123**, 66–85 (2018).
31. Dunn, F. E. et al. Projections of declining fluvial sediment delivery to major deltas worldwide in response to climate change and anthropogenic stress. *Environ. Res. Lett.* **14**, 084034 (2019).
32. Anthony, E. J. et al. Linking rapid erosion of the Mekong River delta to human activities. *Sci. Rep.* **5**, 14745 (2015).
33. Minderhoud, P. S. J. et al. The relation between land use and subsidence in the Vietnamese Mekong delta. *Sci. Total Environ.* **634**, 715–726 (2018).
34. Keogh, M. E. & Törnqvist, T. E. Measuring rates of present-day relative sea-level rise in low-elevation coastal zones: a critical evaluation. *Ocean Sci.* **15**, 61–73 (2019).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2020

## Methods

We predict delta morphology and delta morphologic change by calculating potential sediment transport fluxes due to waves, tides and the river. We obtain delta land area change by summing land gain and land loss from recent global surface-water change studies<sup>20,35</sup>. Our method involves the following seven steps, including estimates of uncertainty: (1) locating coastal river deltas globally, (2) obtaining the pristine and disturbed fluvial sediment flux for each delta, (3) calculating the wave-driven and (4) the tide-driven sediment flux for each delta, (5) producing a morphological prediction for each delta, (6) testing the morphological prediction and (7) obtaining rates of delta land area change.

### Locating river deltas

We locate coastal deltas using HydroSheds at a resolution of 15 arcsec for all coasts south of 60° latitude<sup>36</sup>. HydroSheds uses hydrologically conditioned Shuttle Radar Topography Mission (SRTM)<sup>37</sup> data to generate gridded hydrologic data such as drainage direction and flow accumulation, and includes locations of river mouths globally.

The 15-arcsec HydroSheds dataset contains about 2.48 million first-order drainage basins; 85% of those are smaller than 1 km<sup>2</sup> (ref.<sup>38</sup>). Most of these small drainage basins have no river<sup>38</sup>, and therefore also no delta. They appear mostly along coastlines because of elevation noise that leads to poor drainage delineation of flat, low-lying areas<sup>39</sup> (Extended Data Fig. 1). For studies that focus on rivers, a common solution to this problem is to limit the analysis to drainage basins larger than a certain size (for example, 40,000 km<sup>2</sup>)<sup>14</sup>. Unfortunately, this solution is not appropriate for our purposes because it would exclude many of the smaller deltas. Instead, we select river mouths with a drainage area of at least 50 km<sup>2</sup> if it contains a drainage divide higher than 40 m above mean sea level. We also include drainage basins larger than 1,000 km<sup>2</sup> regardless of the drainage basin topography. Accounting for drainage area elevations in small basins allows us to exclude most of the coastal noise caused, for example, by vegetation, but still captures many small, mountainous drainage basins. We find drainage divide elevations for all river mouths from our initial selection by extracting the SRTM elevation along each drainage basin boundary (Extended Data Fig. 1).

For latitudes greater than 60°, where HydroSheds is not available, we find deltas by selecting drainage basins larger than 1,000 km<sup>2</sup> based on the 1-min ETOPO1<sup>40</sup> grid, which is available globally. We eliminate non-coastal deltas by only selecting potential delta-mouth locations closer than 12 arcmin to the National Oceanic and Atmospheric Administration (NOAA) shoreline (~15 km, depending on the latitude)<sup>41</sup>.

To further improve our dataset and include only alluvial river mouths, we use the WBMSed 2.0 distributed global-scale sediment flux model<sup>14,42</sup> and retrieve river discharge and sediment flux for each river mouth (see Methods section 'Fluvial sediment flux  $Q_{\text{river}}$ '). We remove river mouths with a river discharge below 1 m<sup>3</sup> s<sup>-1</sup> or a sediment flux below 0.01 kg s<sup>-1</sup> (arid environments). We use the global coastal typology dataset of Dürr et al.<sup>43</sup> to further remove drainage basins smaller than 1,000 km<sup>2</sup> that drain into fjords, where  $R$  and  $T$  are unlikely to be appropriate indicators of their morphology. Our resulting dataset consists of 10,848 deltas on all major landmasses except Antarctica and Greenland.

We investigate whether our criteria lead to the inclusion of most coastal deltas globally by creating a test dataset of deltas on Madagascar. Madagascar has a wide range of wave exposure, tidal amplitudes and, consequently, coastal environments. Using 1-m-resolution DigitalGlobe images we visually identify 306 river mouths, of which 236 appear deltaic (where the coastal morphology is affected by the presence of a river; see .kml file at <https://doi.org/10.17605/OSF.IO/S28QB>). Of the 236 deltas, our algorithm finds 212, and 24 deltas were

not located (false negatives, generally small deltas). Our dataset also includes 12 drainage basins that do not have a delta (false positives); these tend to be tributaries to other rivers with confluences near the coast, or small drainage basins without an observable river. We include bayhead deltas in our dataset.

Our test dataset allows us to compute the uncertainty on the global number of deltas (Extended Data Table 1). Combined, our assessment indicates an accuracy of 85%. By extrapolating globally outside Madagascar and following Olofsson et al.<sup>44</sup>, we obtain a standard deviation of 252 and 95% confidence bounds of  $\pm 494$ . Because our false-negative and false-positive rates are comparable, our estimate of 10,848 coastal deltas is unlikely to be strongly biased<sup>44</sup>.

### Fluvial sediment flux $Q_{\text{river}}$

To estimate the fluvial sediment flux for every delta, we use the WBMSed 2.0 distributed global-scale sediment flux model<sup>14,42</sup>. WBMSed is an empirical model that calculates gridded daily fluvial water discharge on the basis of precipitation, temperature, soil type, elevation and other datasets, in this case for the years 1980–2010. Sediment discharge is then estimated using the BQART model<sup>45</sup>.

WBMSed is available globally at a resolution of 6 arcmin, which is lower than that of the HydroSheds data. We therefore convert the WBMSed accumulated discharge and sediment flux file to a discharge and sediment yield (Extended Data Fig. 2). We then sum the discharge and sediment yield across the drainage basins to calculate a discharge and  $Q_{\text{river}}$  for each delta.

WBMSed accounts for human influences on fluvial sediment fluxes by including empirically tested trapping coefficients for river dams and human erosion parameters to account for land-use changes. By disabling these coefficients, WBMSed can estimate fluvial sediment fluxes for a world without humans<sup>42</sup>. We use 'pristine' (without humans) and 'disturbed' (with humans) model results from Cohen et al.<sup>42</sup> to investigate human-induced changes to delta morphology (Extended Data Fig. 3). We note that depending on the history of anthropogenic change, pristine conditions can refer to different time periods, depending on the drainage basin. The Mekong River Delta, for example, has had a long history of human impact on its fluvial sediment flux<sup>46</sup>. Disturbed conditions refer to the present day and include the effects of afforestation and improved soil conservation practices on the fluvial sediment flux to river deltas. WBMSed is validated by independent measurements of the fluvial sediment flux of pristine and disturbed drainage basins<sup>42</sup>. We note that both realizations are based on the 1980–2010 hydroclimate, so we exclude the effects of longer-term climate change on the fluvial sediment flux.

WBMSed provides a reasonable prediction of sediment discharge as tested against observations ( $R^2 = 0.66$ )<sup>14</sup>. Sediment flux estimates remain challenging; therefore, predictions might differ from local case studies, both for pristine and for disturbed river basin conditions. WBMSed data should be considered estimates.

### Wave sediment flux $Q_{\text{wave}}$

To assess ocean wave effects on delta morphology, we calculate the maximum potential alongshore sediment flux  $Q_{\text{wave}}$  (ref.<sup>12</sup>) for every delta using the NOAA WaveWatch III 30-year hindcast phase II<sup>47</sup> by extracting the angular distribution of the wave energy, the significant wave height and the wave period (Extended Data Fig. 4). The resolution of the wave data varies between 4 arcmin and 30 arcmin depending on location and bathymetric complexity. We extract the closest available wave data for each delta.

We calculate  $Q_{\text{wave}}$  by convolving the angular distribution of wave energy with an approximation of alongshore sediment transport recasted into deep-water wave properties

$$Q_{\text{wave}} = \max_{-\pi \leq \theta \leq \pi} [E(\varphi_0) Q_s(\varphi_0 - \theta)] - \min_{-\pi \leq \theta \leq \pi} [E(\varphi_0) Q_s(\varphi_0 - \theta)] \quad (1)$$

# Article

where  $E$  (dimensionless) is the relative contribution of each wave approach angle  $\phi_0$  to alongshore sediment transport.  $Q_s$  (in kilograms per second) represents wave-driven alongshore sediment transport posed in deep-water terms as a function of the approach angle of the wave,  $\phi_0$ , compared to the shoreline  $\theta$  (refs. <sup>12,48</sup>). We do not have global data of shoreline orientation, and therefore calculate  $Q_{\text{wave}}$  by assuming maximum potential transport to the left and the right, away from the river mouth<sup>12</sup>. Given that most of the wave energy is directed towards the coast (not away from the coast), this is unlikely to be a major component of the uncertainty.

Our analysis assumes that waves refract and shoal over shore-parallel contours<sup>12,48</sup> and that the delta is exposed to waves from all directions. Complex nearshore bathymetry and shadowing by headlands can have a considerable effect on wave transformations, but cannot be accounted for in this global model. We therefore assume that if wave data are found within 1° of the river mouth, the delta is not sheltered from wave attack. We assume negligible wave-driven sediment transport if the delta is located farther than 1° from available wave data (sheltered, mostly bayhead deltas). This cutoff could falsely identify some bayhead deltas as wave-dominated, whereas other open-coast deltas might be labelled river-dominated owing to the coarse WaveWatch III grid resolution. We note that this is an important simplification that should be improved upon in the future.

The fluvial dominance ratio  $R$  compares the wave-driven flux  $Q_{\text{wave}}$  to the fluvial sediment that is retained nearshore. WBMSED predicts fluvial suspended load sediment fluxes, of which a large fraction will probably be lost to the marine environment. Bedload fluxes are more likely to be retained nearshore, but no global data exist to predict these fluxes. Here we assume that WBMSED approximates the fluvial sediment load that is retained nearshore. This assumption will most probably lead to an underestimation of wave dominance for larger, suspended-load dominated rivers and an overestimation of wave dominance for smaller, bedload dominated rivers.

The fluvial dominance ratio  $R$  is dependent on the number of distributary channels. The potential alongshore transport  $Q_{\text{wave}}$  acts on each river mouth, whereas  $Q_{\text{river}}$  is split between river mouths<sup>12</sup>. Because no global data on distributary channel networks exist we neglect the effect of distributary formation on  $Q_{\text{wave}}$ , and therefore might underpredict wave influence on deltas with multiple distributaries (for example, Mekong Delta<sup>49</sup>).

## Tidal sediment flux $Q_{\text{tide}}$

We calculate  $Q_{\text{tide}}$  for every coastal delta to establish the effect of tides on delta morphology.  $Q_{\text{tide}}$  is a tidal sediment flux amplitude at the mouth of a delta. If  $Q_{\text{tide}}$  is large compared to  $Q_{\text{river}}$ , we predict considerable channel widening compared to the upstream (fluvial) channel width.  $Q_{\text{tide}}$  requires estimates of the tidal amplitude, angular frequency, channel cross-sectional aspect ratio and channel slope<sup>13</sup>. We extract the tidal amplitude and angular frequency of 13 tidal constituents globally for all deltas using the 15-arcsec-resolution OSU TOPEX dataset<sup>50</sup> (Extended Data Fig. 5). We define the mean tidal amplitude as half of the sum of all tidal constituents and use either a semi-diurnal or a diurnal frequency, depending on the delta location.

We estimate the channel slope from the HydroSheds accumulated drainage area data (ACC files)<sup>36</sup> and the global SRTM data<sup>37</sup> by tracking the elevation upstream from every delta up to 20 m above the mean sea level (Extended Data Fig. 1b). We then fit an exponential function to the elevation data and calculate the gradient of that function at sea level<sup>13</sup>. We assume a slope of  $1 \times 10^{-3}$  (median slope of all coastal deltas) if SRTM elevation data are missing ( $>60^\circ$  latitude) or if its resolution is insufficient to capture the water-surface elevation of deltas.

Nienhuis et al.<sup>13</sup> defined tidal dominance as the ratio of tidal discharge amplitude ( $Q_{\text{w,tide}}$ , in cubic metres per second) and the mean annual river discharge ( $Q_{\text{w,river}}$ , in cubic metres per second). To compare tidal dominance to wave dominance, here we define an equivalent tidal

sediment flux  $Q_{\text{tide}}$  by assuming that the sediment concentration of the tidal discharge is equal to the sediment concentration of the river discharge. We estimate  $Q_{\text{tide}}$  as

$$Q_{\text{tide}} = Q_{\text{w,tide}} \frac{Q_{\text{river}}}{Q_{\text{w,river}}} \quad (2)$$

such that the ratio  $T$  in discharge terms is equivalent to the ratio posed in sediment fluxes. We calculate  $Q_{\text{w,tide}}$  by

$$Q_{\text{w,tide}} = \frac{1}{2} \omega k a^2 \left( \frac{d_u}{S} \right)^2 \beta \quad (3)$$

where  $\omega$  is the tidal angular velocity ( $\text{s}^{-1}$ );  $k$  is a proportionality coefficient ( $\text{m}^{-1}$ ) that is dependent on the grain size, Shields stress and flow roughness<sup>13</sup>;  $a$  is the mean tidal amplitude (m) (Extended Data Fig. 5);  $d_u$  is the upstream channel depth (m);  $S$  is the channel slope; and  $\beta$  is the channel aspect ratio. We estimate the aspect ratio and depth of each river based on its discharge following hydraulic geometry<sup>16</sup>.  $Q_{\text{tide}}$  has been tested for a broad selection of deltas globally and was found to be an appropriate indicator of tidal dominance in a broad range of wave environments<sup>13</sup>.

## Combining $Q_{\text{river}}$ , $Q_{\text{tide}}$ and $Q_{\text{wave}}$

To estimate the location of deltas within a ternary diagram we determine the fraction  $r$  of the total sediment flux contributed by waves, tides and the river

$$r_x = \frac{Q_x}{Q_{\text{river}} + Q_{\text{wave}} + Q_{\text{tide}}} \quad (4)$$

where  $x$  represents river, wave or tide. The relative sediment flux  $r$  can vary between 0 and 1, whereas the river- and tidal-dominance ratios  $R$  and  $T$  vary between  $1/\infty$  and  $\infty$  (Fig. 1a, b).  $r$  allows us to uniquely position a river delta within the ternary diagram and characterize its two first-order morphological indicators, the delta protrusion angle and the channel width divergence. Similarly to wave, tide and river dominance, a delta is considered tide-dominated if  $Q_{\text{tide}}$  exceeds both  $Q_{\text{river}}$  and  $Q_{\text{wave}}$ . By assessing  $Q_{\text{river}}$ ,  $Q_{\text{tide}}$  and  $Q_{\text{wave}}$  for all deltas globally, we find that 8,551 (79%) are wave-dominated, 1,170 (11%) are river-dominated and 1,127 (10%) are tide-dominated.

## Accuracy of delta morphology prediction

To test our predictions of delta morphology, we analysed 212 deltas on Madagascar, supplemented by 100 deltas picked randomly from our dataset, and visually categorized them as river-, wave- or tide-dominated (Extended Data Table 2). Following Olofsson et al.<sup>44</sup>, we obtain prediction accuracies of 91%, 55% and 64%, for wave-, river- and tide-dominated deltas, respectively, which indicate the likelihood that any one particular delta is classified correctly (equation 2 in ref. <sup>44</sup>). By weighting by their occurrence, we obtain an overall accuracy of 85% ( $\pm 2\%$ , determined through bootstrapping) (equation 4 in ref. <sup>44</sup>). By correcting for the size of the dataset, we obtain estimates of the 95% confidence interval of the global fraction of wave-, river- and tide-dominated deltas of  $79\% \pm 9\%$ ,  $11\% \pm 2\%$ , and  $10\% \pm 3\%$ , respectively (equation 11 in ref. <sup>44</sup>).

We note that although the island of Madagascar has a large variety of coastal landforms, it is not necessarily a good statistical representation of coastlines worldwide. Our morphological accuracy assessment is therefore biased, and we do not adjust the gross total proportion of river-, wave- or tide-dominated deltas on the basis of our visual assessment.

## Measurements of recent deltaic change

We measure the deltaic surface area change by combining our dataset of river mouths and their associated deltas with surface-water changes

between 1985 and 2015 mapped on a global scale by the Aqua Monitor<sup>20</sup>. To select the appropriate coastal change per delta we first determine delta extents along the NOAA vectorized shoreline dataset<sup>41</sup>. Next, we use an empirical approximation of the delta area<sup>51</sup>,  $\sim 1.07(Q_{\text{river}}^{1.1} Q_{\text{w,river}}^{0.45})/D_{\text{sh}}$  (in square kilometres), where  $Q_{\text{w,river}}$  is the river discharge and  $D_{\text{sh}}$  is the shelf depth, here  $D_{\text{sh}} \approx 100$  m (ref.<sup>51</sup>). We obtain a delta radius ( $\sim (\text{area}/\pi)^{1/2}$ ), set a minimum radius of 2 km for small deltas, and match every shoreline location within the radius of that particular delta (Extended Data Fig. 5). Using Google Earth Engine<sup>52</sup>, we then retrieve local surface-water changes along these deltaic coastlines, summing land gain and land loss along the NOAA vectorized shorelines within a buffer equal to one-tenth of the delta radius (Extended Data Fig. 5). The NOAA shorelines include banks of wide coastal channels such as estuaries. By selecting only land area change near the NOAA shorelines, we exclude land–water conversion within delta interiors (away from channel banks and shorelines), for which  $R$  and  $T$  are not appropriate indicators. Land area change resulting from, for example, subsidence, tectonic activity, or delta plain engineering, is therefore probably not fully captured in our reported delta- area change. Land area change of abandoned delta lobes near active parts of the delta might be included. We note the potential for sizeable anthropogenic effects on land gain and land loss (for example, land reclamation), and therefore mask out portions of each delta that are classified as urban/artificial (class 190) areas by the GlobCover<sup>53</sup> dataset.

We estimate the uncertainty in the land gain and land loss measurements by combining three sources of error. The first source of error lies in the per-pixel classification of water versus land. The Global Surface Water Explorer reports uncertainty of about 1% in their classification<sup>35</sup>. The Aqua Monitor uses a similar classification algorithm and therefore probably has similar uncertainty. The second source of error is the categorization of changes in the water-to-land and land-to-water transition. We estimate this uncertainty by comparing deltaic land area changes between the AquaMonitor<sup>20</sup> and the Global Surface Water Explorer<sup>35</sup>, which use different algorithms to classify transitions. We obtain a covariance of 7%, which we include as a measure of the spatial uncertainty.

A third source of uncertainty is the shoreline length and buffer assigned to every delta, and how much of the change within and outside that area is related to delta morphodynamics. To quantify this uncertainty, we manually map the surface extents of 40 deltas in Madagascar and measure land surface changes within those deltas. A comparison with automatically mapped areas yields a standard error of 1%. We combine the three sources of uncertainty and obtain a standard error of the mean of 9% per delta. The total net deltaic land area change  $\pm 2$  s.d. for the 10,848 deltas in the dataset between 1985 and 2015 is  $54 \pm 12$  km<sup>2</sup>.

Aside from a global assessment, we also compare land gain rates of specific deltas to values reported by case studies in the literature (Extended Data Table 5). For the Mississippi Delta comparison, we therefore include land loss rates of the ‘birdfoot’ area closest to the river mouth, as well as the Breton Sound basin as defined by Couvillion et al.<sup>22</sup>. For the seven deltas considered, the global analysis seems to capture delta land loss and land gain in the same order of magnitude. Because the time periods and spatial coverages of these studies do not align, we use this only to illustrate similarities and differences between our reported land gain and earlier studies.

## Data availability

All primary sources (OSU TOPEX<sup>50</sup>, NOAA WaveWatch<sup>47</sup>, USGS HydroSheds<sup>36</sup>, USGS SRTM<sup>37</sup>, WBMSed<sup>42</sup> and AquaMonitor<sup>20</sup> data) are publicly available. Wave and tide data can also be found at <https://jhnienhuis.users.earthengine.app>. The resulting morphological predictions for all 10,484 deltas are available as .mat and .kml files at <https://doi.org/10.17605/OSF.IO/S28QB>. Source data for Figs. 1–3 are provided with the paper.

## Code availability

The Matlab computer code that reproduces our findings is available at <https://github.com/jhnienhuis/GlobalDeltaChange> and <https://osf.io/s28qb/>.

35. Pekel, J.-F., Cottam, A., Gorelick, N. & Belward, A. S. High-resolution mapping of global surface water and its long-term changes. *Nature* **540**, 418–422 (2016).
36. Lehner, B., Verdin, K. & Jarvis, A. New global hydrography derived from spaceborne elevation data. *Eos* **89**, 93–94 (2008).
37. USGS. Shuttle Radar Topography Mission. *Global Land Cover Facility*, University of Maryland, College Park, Maryland (USGS, 2006); <https://www2.jpl.nasa.gov/srtm/>.
38. Guth, P. L. Drainage basin morphometry: a global snapshot from the shuttle radar topography mission. *Hydrol. Earth Syst. Sci.* **15**, 2091–2099 (2011).
39. Rahman, M. M., Arya, D. S. & Goel, N. K. Limitation of 90 m SRTM DEM in drainage network delineation using D8 method—a case study in flat terrain of Bangladesh. *Appl. Geomatics* **2**, 49–58 (2010).
40. Amante, C. & Eakins, B. W. *ETOPO1 1 Arc-Minute Global Relief Model: Procedures, Data Sources and Analysis*. NOAA Technical Memorandum NESDIS NGDC-24 (NOAA, 2009).
41. NOAA. NOAA GSHHG Coastline (National Geophysical Data Center, 2015); <http://www.ngdc.noaa.gov/mgg/shorelines/shorelines.html>.
42. Cohen, S., Kettner, A. J., Syvitski, J. P. M. & Fekete, B. M. WBMSed, a distributed global-scale riverine sediment flux model: model description and validation. *Comput. Geosci.* **53**, 80–93 (2013).
43. Dürr, H. H. et al. Worldwide typology of nearshore coastal systems: defining the estuarine filter of river inputs to the oceans. *Estuaries Coasts* **34**, 441–458 (2011).
44. Olofsson, P. et al. Good practices for estimating area and assessing accuracy of land change. *Remote Sens. Environ.* **148**, 42–57 (2014).
45. Kettner, A. J. & Syvitski, J. P. M. HydroTrend v.3.0: a climate-driven hydrological transport model that simulates discharge and sediment load leaving a river system. *Comput. Geosci.* **34**, 1170–1183 (2008).
46. Wang, H. et al. Recent changes of sediment flux to the western Pacific Ocean from major rivers in East and Southeast Asia. *Earth Sci. Rev.* **108**, 80–100 (2011).
47. Chavla, A., Spindler, D. M. & Tolman, H. L. Validation of a thirty year wave hindcast using the Climate Forecast System Reanalysis winds. *Ocean Model.* **70**, 189–206 (2013).
48. Ashton, A. D. & Murray, A. B. High-angle wave instability and emergent shoreline shapes: 2. Wave climate analysis and comparisons to nature. *J. Geophys. Res.* **111**, F04012 (2006).
49. Tamura, T. et al. Origin and evolution of inter-distributary delta plains: insights from Mekong River delta. *Geology* **40**, 303–306 (2012).
50. Egbert, G. D. & Erofeeva, S. Y. Efficient inverse modeling of barotropic ocean tides. *J. Atmos. Ocean. Technol.* **19**, 183–204 (2002).
51. Syvitski, J. P. M. & Saito, Y. Morphodynamics of deltas under the influence of humans. *Global Planet. Change* **57**, 261–282 (2007).
52. Gorelick, N. et al. Google Earth Engine: planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* **202**, 18–27 (2017).
53. Bontemps, S., Defourny, P., Bogaert, E., Van, K., Kalogerou, V. & Perez, J. R. *GLOBCOVER 2009 Products Description and Validation Report* (ESA & UCL, 2011).
54. Jiménez, J. A. & Sánchez-Arcilla, A. Medium-term coastal response at the Ebro delta, Spain. *Mar. Geol.* **114**, 105–118 (1993).
55. Hassan, S. M. T., Syed, M. A. & Mamnun, N. Estimating erosion and accretion in the coast of Ganges-Brahmaputra-Meghna Delta in Bangladesh. In *6th Int. Conference on Water & Flood Management* 115–124 (Institute of Water and Flood Management, 2017).
56. Sarwar, M. G. M. & Woodroffe, C. D. Rates of shoreline change along the coast of Bangladesh. *J. Coast. Conserv.* **17**, 515–526 (2013).
57. Kong, D. et al. Evolution of the Yellow River Delta and its relationship with runoff and sediment load from 1983 to 2011. *J. Hydrol.* **520**, 157–167 (2015).
58. Ali, E. M. & El-Magd, I. A. Impact of human interventions and coastal processes along the Nile Delta coast, Egypt during the past twenty-five years. *J. Aquat. Res.* **42**, 1–10 (2016).
59. Besset, M., Anthony, E. J. & Bouchette, F. Multi-decadal variations in delta shorelines and their relationship to river sediment supply: an assessment and review. *Earth Sci. Rev.* **193**, 199–219 (2019).

**Acknowledgements** This research was supported by US National Science Foundation award EAR-1810855, Netherlands Organisation for Scientific Research (NWO) vi.veni.192.123 and a scholarship from the Wageningen University Postdoc Talent Program, all to J.H.N. J.C.R.’s efforts were supported by the DOE BER Regional & Global Climate Modeling Program through the HiLAT project. D.A.E. was supported by National Science Foundation awards 1812019 and 1426997. A.J.F.H. was funded by the NWO within Vici project ‘Deltas out of shape: regime changes of sediment dynamics in tide-influenced deltas’ (grant NWO-TTW 17062). We thank P.J.J.F. Torfs (Wageningen University and Research) for help with the adopted statistical methodology.

**Author contributions** J.H.N., A.D.A. and D.A.E. conceived the study. A.J.K. assisted with the global sediment flux calculations. J.H.N. carried out the study and wrote the initial draft. J.H.N., A.J.F.H. and T.E.T. discussed the results. All authors contributed to the writing of the manuscript.

**Competing interests** The authors declare no competing interests.

## Additional information

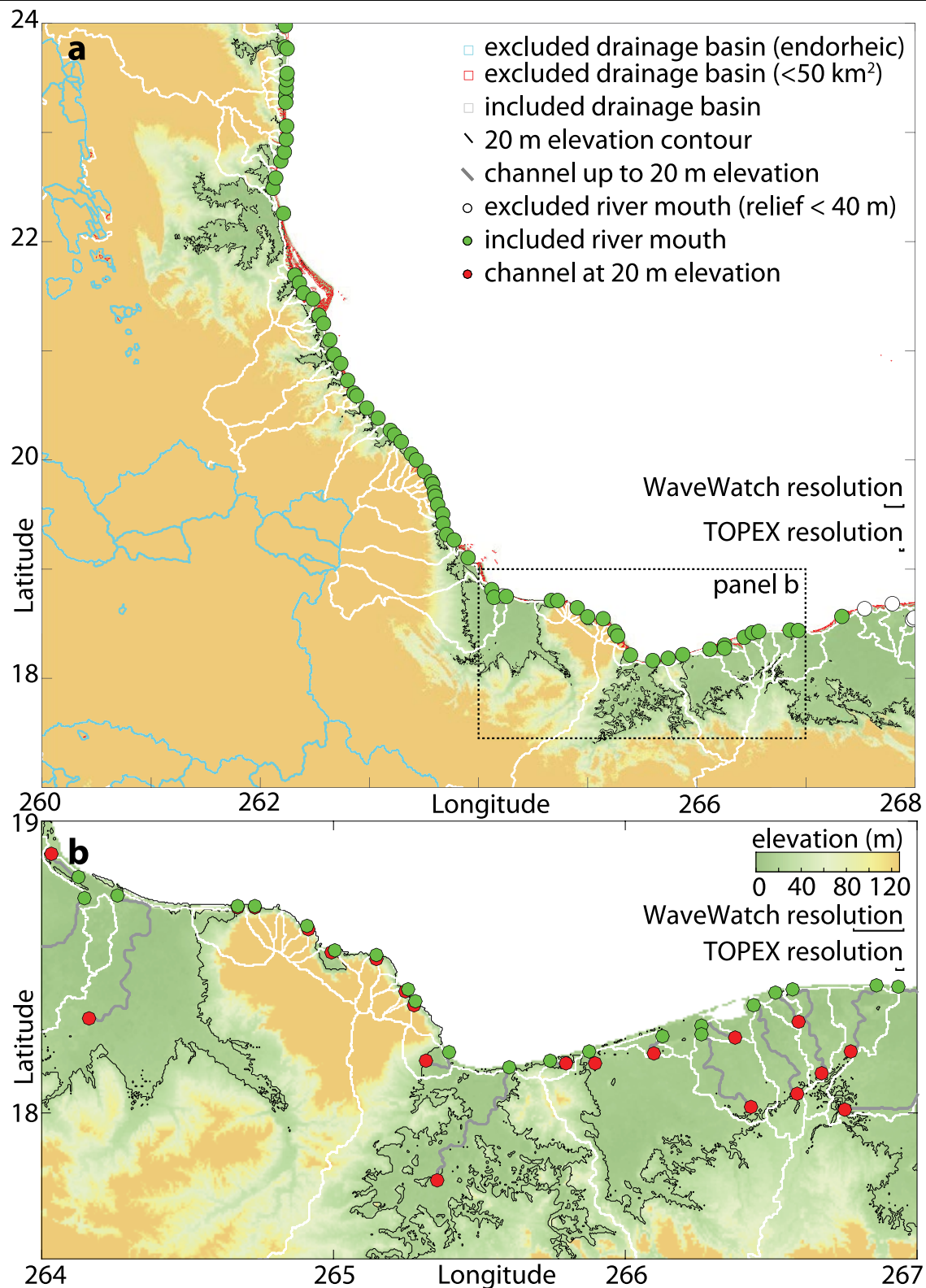
**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41586-019-1905-9>.

**Correspondence and requests for materials** should be addressed to J.H.N.

**Peer review information** *Nature* thanks Nick van de Giesen and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

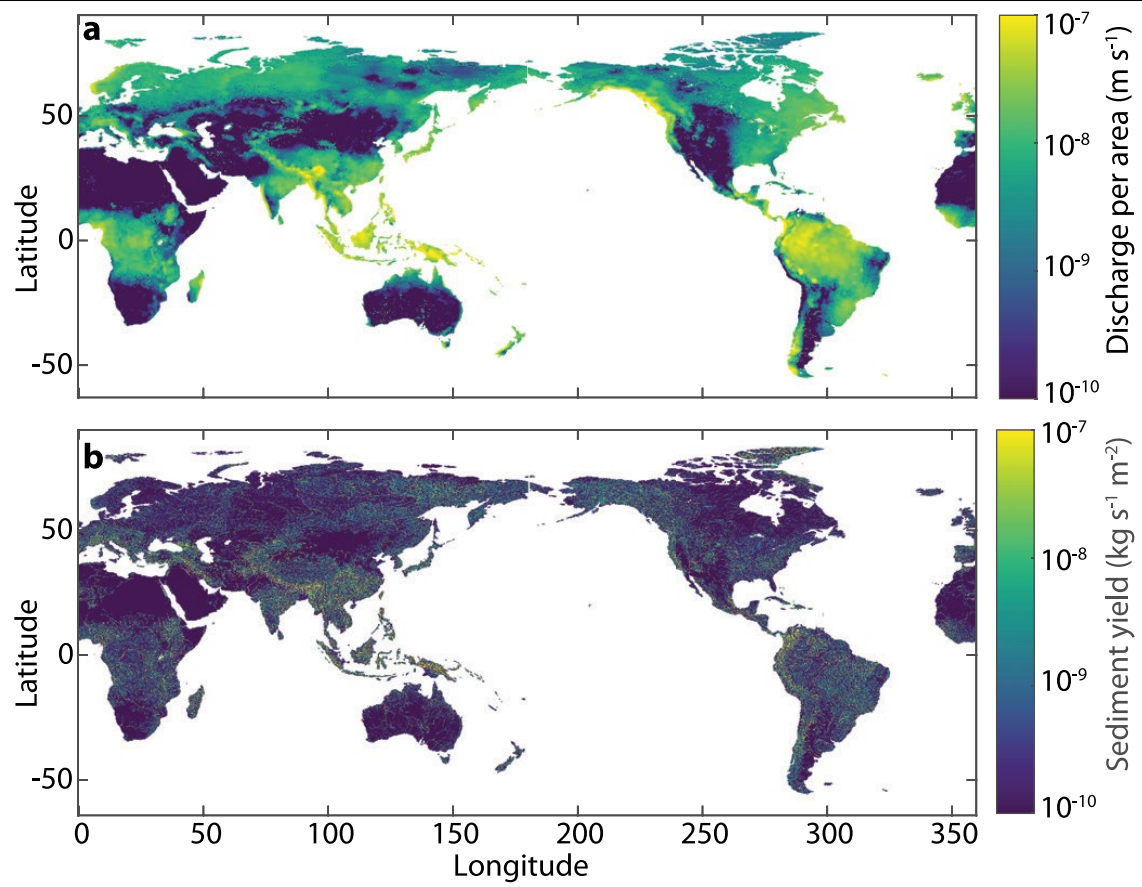
**Reprints and permissions information** is available at <http://www.nature.com/reprints>.



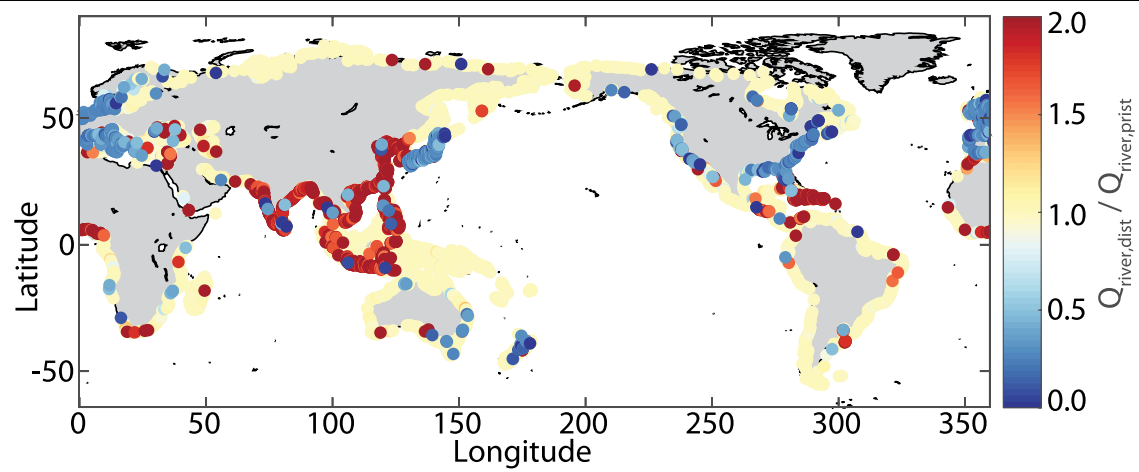


**Extended Data Fig. 1 | Overview of the algorithm that identifies river deltas using HydroSheds data. a.** HydroSheds drainage basins and the included deltas are shown for Veracruz, Mexico. **b.** Close-up of **a**, showing the included

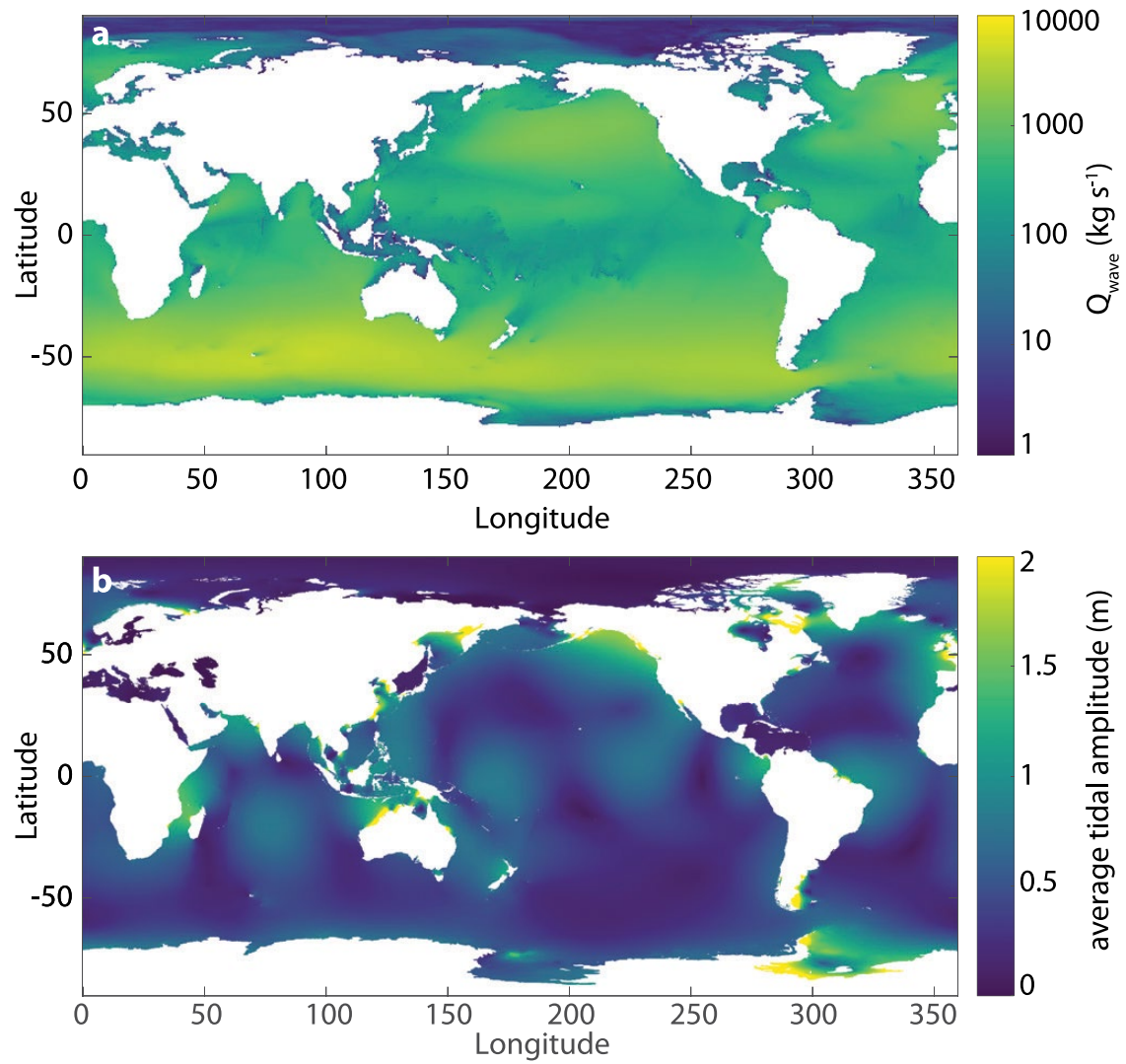
deltas and the tracked river channel for the channel slope calculation. Scale bars show the resolution of the WaveWatch<sup>47</sup> and TOPEX datasets<sup>50</sup>.



**Extended Data Fig. 2 | WBMSed model predictions. a, Discharge per cell. b, Sediment yield<sup>42</sup>.**

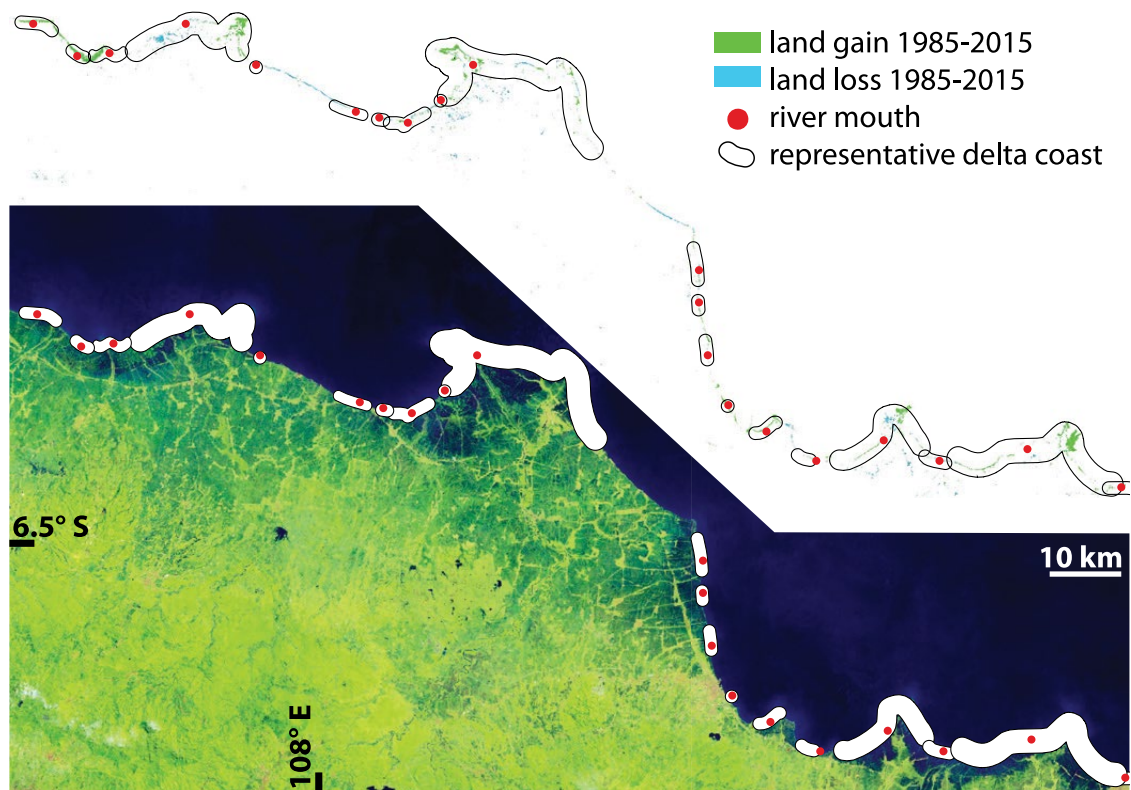


**Extended Data Fig. 3 | WBMsed model predictions of human-induced change to the deltaic fluvial sediment flux.** Colours indicate the ratio of the modern fluvial sediment flux ( $Q_{river}^d$ ; here  $Q_{river,dist}$ ) to the flux in a world without anthropogenic modifications<sup>42</sup> ( $Q_{river}^p$ ; here  $Q_{river,prist}$ ).



**Extended Data Fig. 4 | Characterization of data used for wave- and tide-driven deltaic sediment flux. a**, Global maximum potential alongshore sediment transport ( $Q_{\text{wave}}$ ) based on the WaveWatch 30-year hindcast data<sup>47</sup>. **b**, Global estimate of mean tidal amplitude based on the OSU TOPEX data<sup>30</sup>.





**Extended Data Fig. 5 | Example of recent deltaic land area change for the north shore of Java, Indonesia.** Land loss and land gain were measured using Landsat (<http://landsat.usgs.gov/>) images from Google Earth Engine<sup>52</sup> based on the Deltares Aqua Monitor<sup>35</sup>. Here, deltas have expanded recently because

of human-induced increases in the fluvial sediment flux. The top image shows the coastal change, with the red markers and black outlines representing individual deltas and their coastlines, respectively.

Extended Data Table 1 | Confusion matrix of the number of deltas on Madagascar

		Observed	
		Delta	No delta
Predicted	Delta	212	12
	No delta	24	-

We note that the true-negative rate (no delta observed, no delta predicted) is infinite and therefore not included in our analysis.

Extended Data Table 2 | Confusion matrix of the delta morphologic prediction based on a validation dataset of 312 deltas

		Observed (% of total)			
		Wave	River	Tide	Predicted total
Predicted (% of total)	Wave	72	3	8	83
	River	2	4	1	7
	Tide	1	0	9	10
	Observed total	74	8	18	100

**Extended Data Table 3 | Yearly deltaic land gain, loss and net gain for different regions**

	<b>Land gain</b> (km <sup>2</sup> yr <sup>-1</sup> )	<b>Land loss</b> (km <sup>2</sup> yr <sup>-1</sup> )	<b>Net land gain</b> (km <sup>2</sup> yr <sup>-1</sup> )
Global	181 ± 8.3	- 127 ± 8.3	54 ± 11.8
East Africa	6 ± 1.6	- 3 ± 1.6	3 ± 2.3
South Asia	42 ± 1.7	- 32 ± 1.7	10 ± 2.4
West Africa	3 ± 1.3	- 3 ± 1.3	1 ± 1.8
Europe	10 ± 2.5	- 3 ± 2.5	8 ± 3.6
Central America	2 ± 1.8	- 1 ± 1.8	1 ± 2.5
Russia	1 ± 2.2	- 1 ± 2.2	0 ± 3.1
East Asia	34 ± 2.5	- 22 ± 2.5	11 ± 3.6
Northern Africa/Middle-East	5 ± 1.2	- 2 ± 1.2	3 ± 1.8
Eastern North America	6 ± 2.2	- 11 ± 2.2	- 4 ± 3.2
Western North America	2 ± 1.6	- 2 ± 1.6	0 ± 2.3
Oceania	6 ± 3.0	- 5 ± 3.0	1 ± 4.3
Eastern South America	33 ± 2.0	- 17 ± 2.0	16 ± 2.9
Western South America	3 ± 1.2	- 2 ± 1.2	2 ± 1.8
Southeast Asia	27 ± 4.2	- 23 ± 4.2	4 ± 5.9

Values represent averages from 1985 to 2015. Error limits indicate 2 s.d.

Extended Data Table 4 | Predicted sediment transport fluxes for a selection of well-known deltas

Delta	River Water Discharge $Q_{w,river}$ ( $m^3 s^{-1}$ )	Pristine Fluvial Sediment Flux $Q^p_{river}$ ( $kg s^{-1}$ )	Disturbed Fluvial Sediment Flux $Q^d_{river}$ ( $kg s^{-1}$ )	Wave-driven Sediment Flux $Q_{wave}$ ( $kg s^{-1}$ )	Tide-driven Sediment Flux $Q_{tide}$ ( $kg s^{-1}$ )	Net land gain ( $km^2 yr^{-1}$ )
Amazon	2.0E+5	3.8E+4	3.1E+4	2.9E-1	7.7E+5	1.0E+1
Arno	5.7E+1	7.0E+1	1.0E+0	1.7E+2	1.4E-1	2.1E-2
Colorado, MX	6.9E+2	3.8E+3	4.1E+0	2.9E+1	7.0E+3	-2.7E-1
Copper	1.2E+3	2.2E+3	3.4E+2	7.7E+2	2.8E+3	-7.1E-2
Danube	6.4E+3	2.1E+3	6.4E+2	2.3E+1	1.7E+1	3.7E-1
Ebro	1.4E+3	5.8E+2	2.8E+1	3.5E+1	4.0E+0	-3.8E-1
Eel, CA	2.4E+2	5.6E+2	7.5E+1	2.5E+3	1.5E+2	-1.7E-1
Elbe	4.2E+2	4.9E+2	2.5E+2	9.8E+0	4.1E+6	-2.7E-4
Ganges-Brahmaputra	3.1E+4	3.5E+4	3.5E+4	0.0E+0	2.0E+6	4.9E+0
Godavari	2.7E+3	5.4E+3	5.2E+3	3.8E+2	1.0E+2	5.0E-1
Huanghe	1.5E+3	3.5E+4	3.8E+3	2.3E+1	1.8E+1	-8.3E+0
Klamath	4.7E+2	3.2E+2	1.5E+2	2.4E+3	1.9E+3	1.2E-2
Lena	1.6E+4	6.3E+2	5.1E+3	1.2E+0	7.8E+2	7.1E-3
Mekong	1.7E+4	3.1E+3	3.0E+3	3.3E+1	4.0E+5	-2.1E-1
Mississippi	1.5E+4	1.3E+4	4.2E+3	1.0E+3	9.8E+2	-5.2E+0
Niger	6.1E+3	1.3E+3	8.0E+2	6.1E+2	4.8E+3	-4.9E-2
Nile	3.5E+3	3.8E+3	7.6E+1	2.2E+2	2.5E+2	-7.0E-1
Orange	4.4E+2	2.8E+3	3.0E+2	2.9E+3	1.1E+1	2.4E-1
Parana	1.5E+4	2.8E+3	2.5E+3	0.0E+0	9.0E+2	9.4E-1
Po	1.5E+3	5.5E+2	3.0E+2	4.2E+1	4.9E+2	1.2E-1
Rhine-Meuse	2.0E+3	2.0E+3	5.5E+2	1.2E+2	1.7E+4	6.5E-1
Rhone	1.7E+3	1.9E+3	5.6E+2	1.6E+2	7.9E+1	1.7E-1
Sao Francisco	3.6E+3	2.5E+3	1.7E+3	4.4E+3	1.7E+3	2.0E-2
Schelde	9.8E+1	1.9E+1	5.0E+0	6.8E+2	1.2E+0	3.2E-3
Senegal	6.9E+2	5.6E+2	4.3E+2	5.0E+2	8.9E+3	-4.7E-2
Volga	8.2E+3	6.0E+2	1.5E+3	0.0E+0	7.8E-1	3.8E+0
Yangtze	2.8E+4	1.5E+4	9.0E+3	6.1E+1	2.0E+4	-2.7E+0

See also Fig. 2b.



Extended Data Table 5 | Comparison of net land gain estimates with case studies from the literature

Delta	Net land gain $\pm$ 2 s.d. (km <sup>2</sup> yr <sup>-1</sup> , this study)	Net land gain (km <sup>2</sup> yr <sup>-1</sup> , other studies)	Study period	Source	Note
Ebro	-0.4 $\pm$ 0.2	-0.2	1957-1992	54	Based on shoreline transects
Ganges - Brahmaputra	4.9 $\pm$ 0.2	12.3	1973-2016	55	Hatiya and Bhola districts
Ganges - Brahmaputra	4.9 $\pm$ 0.2	0.4	1989-2009	56	Coastal Bangladesh
Huanghe	-8.2 $\pm$ 0.2	-4.0	1999-2011	57	Modern lobe
Mekong	-0.2 $\pm$ 0.2	0.5	2003-2012	32	Delta distributary mouths
Mississippi	-5.2 $\pm$ 0.2	-0.5	1985-2015	22	Birdfoot region
Mississippi	-5.2 $\pm$ 0.2	-15.0	1985-2015	22	Breton Sound basin
Nile	-0.7 $\pm$ 0.2	-0.2	1990-2014	58	
Parana	0.9 $\pm$ 0.2	2.0	1995-2015	59	

Case studies from refs. <sup>22,32,54-59</sup>.

# Isolation of an archaeon at the prokaryote–eukaryote interface

<https://doi.org/10.1038/s41586-019-1916-6>

Received: 6 August 2019

Accepted: 5 December 2019

Published online: 15 January 2020

Open access

Hiroyuki Imachi<sup>1,11\*</sup>, Masaru K. Nobu<sup>2,11\*</sup>, Nozomi Nakahara<sup>1,2,3</sup>, Yuki Morono<sup>4</sup>, Miyuki Ogawara<sup>1</sup>, Yoshihiro Takaki<sup>1</sup>, Yoshinori Takano<sup>5</sup>, Katsuyuki Uematsu<sup>6</sup>, Tetsuro Ikuta<sup>7</sup>, Motoo Ito<sup>4</sup>, Yohei Matsui<sup>8</sup>, Masayuki Miyazaki<sup>1</sup>, Kazuyoshi Murata<sup>9</sup>, Yumi Saito<sup>1</sup>, Sanae Sakai<sup>1</sup>, Chihong Song<sup>9</sup>, Eiji Tasumi<sup>1</sup>, Yuko Yamanaka<sup>1</sup>, Takashi Yamaguchi<sup>3</sup>, Yoichi Kamagata<sup>2</sup>, Hideyuki Tamaki<sup>2</sup> & Ken Takai<sup>1,10</sup>

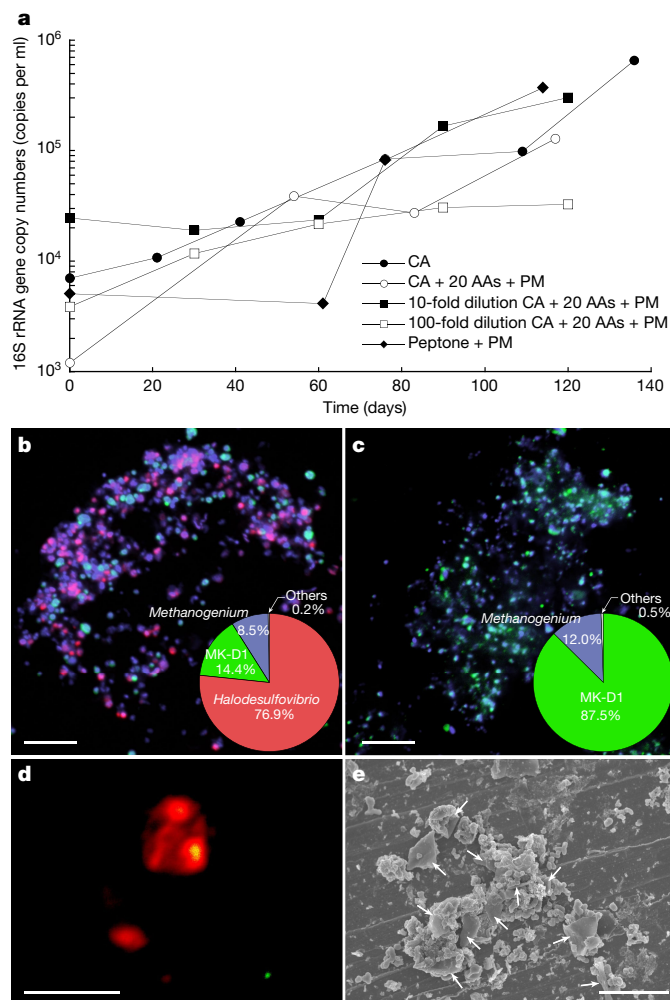
The origin of eukaryotes remains unclear<sup>1–4</sup>. Current data suggest that eukaryotes may have emerged from an archaeal lineage known as ‘Asgard’ archaea<sup>5,6</sup>. Despite the eukaryote-like genomic features that are found in these archaea, the evolutionary transition from archaea to eukaryotes remains unclear, owing to the lack of cultured representatives and corresponding physiological insights. Here we report the decade-long isolation of an Asgard archaeon related to Lokiarchaeota from deep marine sediment. The archaeon—‘*Candidatus Prometheoarchaeum syntrophicum*’ strain MK-D1—is an anaerobic, extremely slow-growing, small coccus (around 550 nm in diameter) that degrades amino acids through syntrophy. Although eukaryote-like intracellular complexes have been proposed for Asgard archaea<sup>6</sup>, the isolate has no visible organelle-like structure. Instead, *Ca. P. syntrophicum* is morphologically complex and has unique protrusions that are long and often branching. On the basis of the available data obtained from cultivation and genomics, and reasoned interpretations of the existing literature, we propose a hypothetical model for eukaryogenesis, termed the entangle–engulf–endogenize (also known as E<sup>3</sup>) model.

How the first eukaryotic cell emerged remains unclear. Among various competing evolutionary models, the most widely accepted are symbiogenic models in which an archaeal host cell and an alphaproteobacterial endosymbiont merged to become the first eukaryotic cell<sup>1–4</sup>. Recent metagenomic characterization of deep-sea archaeal group/marine benthic group-B (also known as Lokiarchaeota) and the Asgard archaea superphylum led to the theory that eukaryotes originated from an archaeon that was closely related to these lineages<sup>5,6</sup>. The genomes of Asgard archaea encode a repertoire of proteins that are only found in Eukarya (eukaryotic signature proteins), including those involved in membrane trafficking, vesicle formation and/or transportation, ubiquitin and cytoskeleton formation<sup>6</sup>. Subsequent metagenomic studies have suggested that Asgard archaea have a wide variety of physiological properties, including hydrogen-dependent anaerobic autotrophy<sup>7</sup>, peptide or short-chain hydrocarbon-dependent organotrophy<sup>8–12</sup> and rhodopsin-based phototrophy<sup>13,14</sup>. However, no representative of the Asgard archaea has been cultivated and, thus, the physiology and cell biology of this clade remains unclear. In an effort to close this knowledge gap, we successfully isolated an archaeon of this clade, report its physiological and genomic characteristics, and propose a new model for eukaryogenesis.

## Isolation of an Asgard archaeon

Setting out to isolate uncultivated deep marine sediment microorganisms, we engineered and operated a methane-fed continuous-flow bioreactor system for more than 2,000 days to enrich such organisms from anaerobic marine methane-seep sediments<sup>15</sup> (Supplementary Note 1). We successfully enriched many phylogenetically diverse yet-to-be cultured microorganisms, including Asgard archaea members (Loki-, Heimdall- and Odinararchaeota)<sup>15</sup>. For further enrichment and isolation, samples of the bioreactor community were inoculated in glass tubes with simple substrates and basal medium. After approximately one year, we found faint cell turbidity in a culture containing casamino acids supplemented with four bacteria-suppressing antibiotics (Supplementary Note 2) that was incubated at 20 °C. Clone library-based small subunit (SSU) rRNA gene analysis revealed a simple community that contained *Halodesulfobivibrio* and a small population of Lokiarchaeota (Extended Data Table 1). In pursuit of this archaeon, which we designated strain MK-D1, we repeated subcultures when MK-D1 reached maximum cell densities as measured by quantitative PCR (qPCR). This approach gradually enriched the archaeon, which has an extremely slow growth rate and low cell yield (Fig. 1a). The culture consistently had a 30–60-day lag phase and required more

<sup>1</sup>Institute for Extra-cutting-edge Science and Technology Avant-garde Research (X-star), Japan Agency for Marine–Earth Science and Technology (JAMSTEC), Yokosuka, Japan. <sup>2</sup>Bioproduction Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Japan. <sup>3</sup>Department of Civil and Environmental Engineering, Nagaoka University of Technology, Nagaoka, Japan. <sup>4</sup>Kochi Institute for Core Sample Research, X-star, JAMSTEC, Nankoku, Japan. <sup>5</sup>Biogeochemistry Program, Research Institute for Marine Resources Utilization, JAMSTEC, Yokosuka, Japan. <sup>6</sup>Department of Marine and Earth Sciences, Marine Work Japan, Yokosuka, Japan. <sup>7</sup>Research Institute for Global Change, JAMSTEC, Yokosuka, Japan. <sup>8</sup>Research Institute for Marine Resources Utilization, JAMSTEC, Yokosuka, Japan. <sup>9</sup>National Institute for Physiological Sciences, Okazaki, Japan. <sup>10</sup>Section for Exploration of Life in Extreme Environments, Exploratory Research Center on Life and Living Systems (ExCELLS), National Institute of Natural Sciences, Okazaki, Japan. <sup>11</sup>These authors contributed equally: Hiroyuki Imachi, Masaru K. Nobu. \*e-mail: imachi@jamstec.go.jp; m.nobu@aist.go.jp



**Fig. 1 | Growth curves and photomicrographs of the cultured Lokiarchaeota strain MK-D1.** **a**, Growth curves of MK-D1 in anaerobic medium supplemented with casamino acids (CA) alone; casamino acids with 20 amino acids (AAs) and powdered milk (PM); or peptone with powdered milk. Results are also shown for cultures fed with 10- and 100-fold dilution of casamino acids, 20 amino acids and powdered milk. **b**, **c**, Fluorescence images of cells from enrichment cultures after 8 (**b**) and 11 (**c**) transfers stained with DAPI (violet) and hybridized with nucleotide probes that target MK-D1 (green) and Bacteria (red). Pie charts show the relative abundance of microbial populations based on SSU rRNA gene-tag sequencing (iTAG) analysis. **d**, A fluorescence image of cells from enrichment cultures after 11 transfers hybridized with nucleotide probes that target MK-D1 (green) and *Methanogenium* (red). The FISH experiments were performed three times with similar results. **e**, SEM image of a highly purified co-culture of MK-D1 and *Methanogenium*. White arrows indicate *Methanogenium* cells. We observed four different co-cultures with *Methanogenium*. Representative of  $n = 40$  recorded images. The detailed iTAG-based community compositions of cultures corresponding to each of the images are shown in Supplementary Table 1. Scale bars, 10  $\mu\text{m}$  (**b**, **c**) and 5  $\mu\text{m}$  (**d**, **e**).

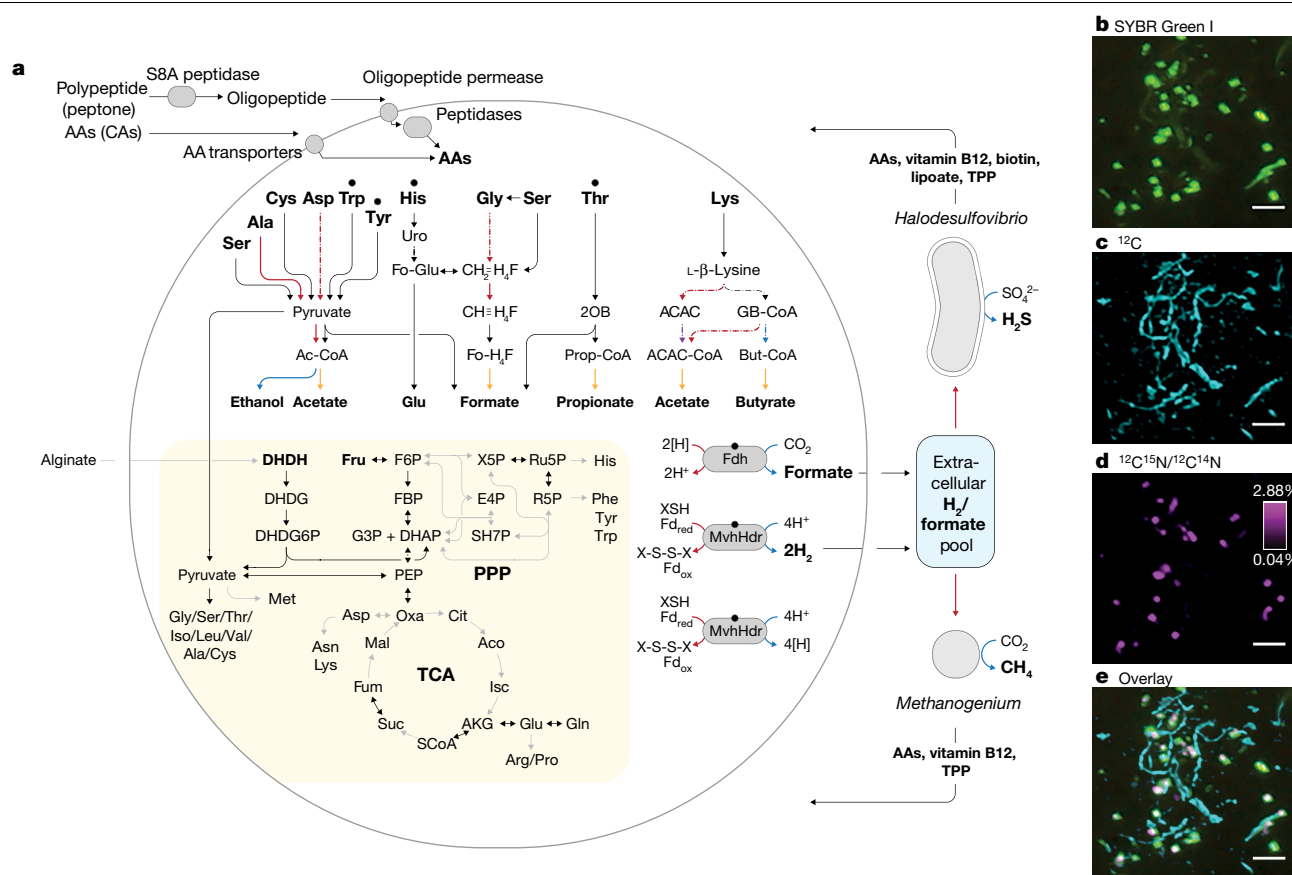
than 3 months to reach full growth: around  $10^5$  16S rRNA gene copies  $\text{ml}^{-1}$  (Fig. 1a). The doubling time was estimated to be approximately 14–25 days. Variation in cultivation temperatures (Extended Data Fig. 1), and substrate combinations and concentrations did not significantly shorten the lag phase or improve growth rate or cell yield (data not shown). Static cultivation supplemented with 20 amino acids and powdered milk resulted in the stable growth. For further characterization, we cultured the archaeon under the optimal conditions determined above.

After six transfers, MK-D1 reached 13% abundance in a tri-culture containing a *Halodesulfobivrio* bacterium (85%) and a *Methanogenium*

archaeon (2%) (Extended Data Table 1). Analyses using fluorescence in situ hybridization (FISH) and scanning electron microscopy (SEM) revealed a close physical association of the archaeon with the other microorganisms (Fig. 1b–e, Extended Data Fig. 3 and Supplementary Table 1). Through metagenome-based exploration of the metabolic potential of this archaeon and a stable-isotope probing experiment, we discovered that MK-D1 can catabolize ten amino acids and peptides through syntrophic growth with *Halodesulfobivrio* and *Methanogenium* through interspecies hydrogen (and/or formate) transfer<sup>16</sup> (Fig. 2, Extended Data Fig. 2 and Supplementary Tables 2–4). Indeed, addition of hydrogen scavenger-inhibiting compounds (that is, 10 mM molybdate and 2-bromoethanesulfonate for sulfate-reducing bacteria (SRB) and methanogens, respectively) significantly impaired growth of MK-D1. Through subsequent transfers, we were able to eliminate the *Halodesulfobivrio* population, enabling us to obtain a pure co-culture of the target archaeon MK-D1 and *Methanogenium* after a 12-year study—from bioreactor-based pre-enrichment of deep-sea sediments to a final 7 years of in vitro enrichment. We here propose the name ‘*Candidatus Prometheoarchaeum syntrophicum*’ strain MK-D1 for the isolated archaeon (see Supplementary Note 3 for reasons why the provisional *Candidatus* status is necessary despite isolation).

## Cell biology, physiology and metabolism

We further characterized MK-D1 using the pure co-cultures and highly purified cultures. Microscopy analyses showed that the cells were small cocci (approximately 300–750 nm in diameter (average, 550 nm)), and generally formed aggregates surrounded by extracellular polymer substances (EPS) (Fig. 3a, b and Extended Data Fig. 3), consistent with previous observations using FISH<sup>15,17</sup>. MK-D1 cells were easily identifiable given the morphological difference from their co-culture partner *Methanogenium* (highly irregular coccoid cells of  $\geq 2 \mu\text{m}$ ; Fig. 1d, e). Dividing cells had less EPS and a ring-like structure around the cells (Fig. 3c). Cryo-electron microscopy (cryo-EM) and transmission electron microscopy (TEM) analyses revealed that the cells contain no visible organelle-like inclusions (Fig. 3d–f and Supplementary Videos 1–6), in contrast to previous suggestions<sup>6</sup>. For cryo-EM, cells were differentiated from vesicles on the basis of the presence of cytosolic material (although DNA and ribosomes could not be differentiated), EPS on the cell surface and cell sizes that were consistent with observations by SEM and TEM analyses (Supplementary Videos 4–6). The cells produce membrane vesicles (50–280 nm in diameter) (Fig. 3b–f) and chains of blebs (Fig. 3c). MK-D1 cells also form membrane-based cytosol-connected protrusions of various lengths that have diameters of 80–100 nm, and display branching with a homogeneous appearance unlike those of other archaea (Fig. 3g–i; confirmed using both SEM and TEM). These protrusions neither form elaborate networks (as in *Pyrodictium*<sup>18</sup>) nor intercellular connections (*Pyrodictium*, *Thermococcus* and *Haloferax*<sup>18–20</sup>), suggesting differences in physiological functions. The MK-D1 cell envelope may be composed of a membrane and a surrounding S-layer, given the presence of four genes that encode putative S-layer proteins (Supplementary Fig. 1), stalk-like structures on the surface of the vesicles (Fig. 3e and Extended Data Fig. 3f, g) and the even distance between the inner and outer layers of the cell envelope (Fig. 3d). Lipid composition analysis of the MK-D1 and *Methanogenium* co-culture revealed typical archaeal isoprenoid signatures— $\text{C}_{20}$ -phytane and  $\text{C}_{40}$ -biphytanes with 0–2 cyclopentane rings were obtained after ether-cleavage treatment (Fig. 3j). Considering the lipid data obtained from a reference *Methanogenium* isolate (99.3% 16S rRNA gene identity; Supplementary Fig. 2), MK-D1 probably contains  $\text{C}_{20}$ -phytane and  $\text{C}_{40}$ -biphytanes with 0–2 rings. The MK-D1 genome encoded most of the genes necessary to synthesize ether-type lipids—although geranylgeranylglycerol phosphate synthase was missing—and lacked genes for ester-type lipid synthesis (Supplementary Tables 3, 4).



**Fig. 2 | Syntrophic amino acid utilization of MK-D1. a**, Genome-based metabolic reconstruction of MK-D1. Metabolic pathways identified (coloured or black) and not identified (grey) are shown. For identified pathways, each step (solid line) or process (dotted) is marked by whether it is oxidative (red), reductive (blue), ATP-yielding (orange) or ATP-consuming (purple). Wavy arrows indicate exchange of compounds: formate, H<sub>2</sub>, amino acids, vitamin B<sub>12</sub>, biotin, lipoate and thiamine pyrophosphate (TPP), which are predicted to be metabolized or synthesized by the partnering *Halodesulfobivibrio* and/or *Methanogenium*. Biosynthetic pathways are indicated with a yellow background. Metatranscriptomics-detected amino-acid-catabolizing pathways are indicated (black dots above amino acids). DHDH, 4,5-dihydroxy-2,6-dioxohexanoate; DHDG, 2-dehydro-3-deoxy-D-gluconate; DHDG6P, 3-dehydro-3-deoxy-D-gluconate 6-phosphate; Ac-CoA, acetyl-CoA; uro, urocanate; Fo-Glu, formyl glutamate; CH<sub>3</sub>=H<sub>4</sub>F, methylene-tetrahydrofolate; CH=H<sub>4</sub>F, methenyl-tetrahydrofolate; Fo-H<sub>4</sub>F, formyl-tetrahydrofolate; 2OB, 2-oxobutyrate; Prop-CoA, propionyl-CoA; ACAC, acetoacetate; GB-CoA,

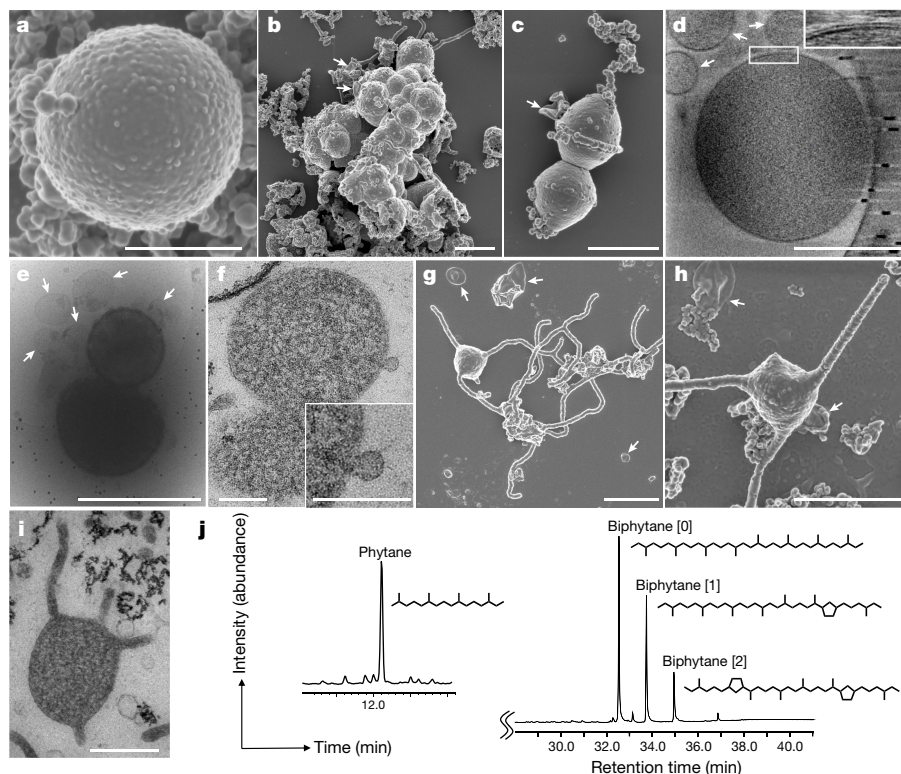
γ-amino-butyryl-CoA; But-CoA, butyryl-CoA; Fd, ferredoxin; XSH/X-S-S-X, thiol/disulfide pair; TCA, tricarboxylic acid cycle; PPP, pentose-phosphate pathway. **b–e**, NanoSIMS analysis of a highly purified MK-D1 culture incubated with a mixture of <sup>13</sup>C- and <sup>15</sup>N-labelled amino acids. **b**, Green fluorescent micrograph of SYBR Green I-stained cells. Aggregates are MK-D1, and filamentous cells are *Methanobacterium* sp. strain MO-MB1 (fluorescence can be weak owing to the high rigidity and low permeability of the cell membrane (Extended Data Fig. 2m, n; see also ref. <sup>49</sup>). **c**, NanoSIMS ion image of <sup>12</sup>C (cyan). **d**, NanoSIMS ion image of <sup>12</sup>C<sup>15</sup>N/<sup>12</sup>C<sup>14</sup>N (magenta). **e**, Overlay image of **b–d**. **d**, The colour bar indicates the relative abundance of <sup>15</sup>N expressed as <sup>15</sup>N/<sup>14</sup>N. Scale bars 5 μm. The NanoSIMS analysis was performed without replicates due to its slow growth rate and low cell density. However, to ensure the reproducibility, we used two different types of highly purified cultures of MK-D1 (see Methods). Representative of *n* = 8 recorded images. The iTAG analysis of the imaged culture is shown in Supplementary Table 1.

MK-D1 can degrade amino acids anaerobically, as confirmed by monitoring the depletion of amino acids during the growth of pure co-cultures (Extended Data Fig. 1b, c). We further verify the utilization of amino acids by quantifying the uptake of a mixture of <sup>13</sup>C- and <sup>15</sup>N-labelled amino acids through nanometre-scale secondary ion mass spectrometry (NanoSIMS) (Fig. 2b–e). Cell aggregates of MK-D1 incorporated amino-acid-derived nitrogen, demonstrating the capacity of MK-D1 to utilize amino acids for growth. Notably, the <sup>13</sup>C-labelling of methane and CO<sub>2</sub> varied depending on the methanogenic partner, indicating that MK-D1 produces both hydrogen and formate from amino acids for interspecies electron transfer (Extended Data Table 2). Indeed, addition of high concentrations of hydrogen or formate completely suppressed growth of MK-D1 (Extended Data Table 3). The syntrophic partner was replaceable—MK-D1 could also grow syntrophically with *Methanobacterium* sp. strain MO-MB1<sup>21</sup> instead of *Methanogenium* (Fig. 2b–e). Although 14 different culture conditions were applied, none enhanced the cell yield, which indicates

specialization of the degradation of amino acids and/or peptides (Extended Data Table 3).

To further characterize the physiology of the archaeon, we analysed the complete MK-D1 genome (Extended Data Fig. 2 and Supplementary Tables 2–6). The genome only encodes one hydrogenase (NiFe hydrogenase MvhADG–HdrABC) and formate dehydrogenase (molybdopter-in-dependent FdhA), suggesting that these enzymes mediate reductive H<sub>2</sub> and formate generation, respectively. MK-D1 represents, to our knowledge, the first cultured archaeon that can produce and syntrophically transfer H<sub>2</sub> and formate using the above enzymes. We also found genes encoding proteins for the degradation of ten amino acids. Most of the identified amino-acid-catabolizing pathways only recover energy through the degradation of a 2-oxoacid intermediate (that is, pyruvate or 2-oxobutyrate; Fig. 2a and Supplementary Table 4). MK-D1 can degrade 2-oxoacids hydrolytically (through 2-oxoacid-formate lyases) or oxidatively (through 2-oxoacid:ferredoxin oxidoreductases) to yield acyl-CoA intermediates that can be further degraded





**Fig. 3 | Microscopy characterization and lipid composition of MK-D1.**

**a–c**, SEM images of MK-D1. Single cell (**a**), aggregated cells covered with EPS-like materials (**b**) and a dividing cell with polar chains of blebs (**c**). **d**, Cryo-electron tomography image of MK-D1. The top-right inset image shows a magnification of the boxed area to show the cell envelope structure. **e**, Cryo-EM image of large membrane vesicles attached to and surrounding MK-D1 cells. **f**, Ultrathin section of an MK-D1 cell and a membrane vesicle. The bottom-right inset image shows a magnified view of the membrane vesicle. **g, h**, SEM images of MK-D1 cells producing long branching (**g**) and straight (**h**) membrane protrusions. **i**, Ultrathin section of a MK-D1 cell with protrusions. **j**, A total ion chromatogram of gas chromatography–mass spectrometry (GC–MS) for lipids extracted from a highly purified MK-D1 culture. The chemical structures of isoprenoids and

their relative compositions are also shown (Supplementary Fig. 2). Scale bars, 1  $\mu\text{m}$  (**b, c, g, h**), 500 nm (**a, d, e, i**) and 200 nm (**f**). **a–c, g, h**, SEM images are representative of  $n = 122$  recorded images that were obtained from four independent observations from four culture samples. **d, e**, Cryo-EM images are representative of  $n = 14$  recorded images that were taken from two independent observations from two culture samples. **f, i**, The ultrathin section images are representative of  $n = 131$  recorded images that were obtained from six independent observations from six culture samples. White arrows in the images indicate large membrane vesicles. The lipid composition experiments were repeated twice and gave similar results. Detailed iTAG-based community compositions of the cultures are shown in Supplementary Table 1.

for ATP generation. In the hydrolytic path, the carboxylate group of the amino acid is released as formate that can be directly handed off to partnering methanogenic archaea or SRB. In the oxidative path, 2-oxoacid oxidation is coupled with release of amino acid carboxylate as  $\text{CO}_2$  and reduction of ferredoxin, which can be re-oxidized through  $\text{H}^+$  and/or  $\text{CO}_2$  reduction to  $\text{H}_2$  and formate, respectively (through the electron-confining NiFe hydrogenase MvhADG–HdrABC or formate dehydrogenase FdhA). On the basis of  $^{13}\text{C}$ -amino-acid-based experiments (Supplementary Note 4), MK-D1 can probably switch between syntrophic interaction through 2-oxoacid hydrolysis and oxidation depending on the partner(s).

**Etymology.** *Prometheoarchaeum*, *Prometheus* (Greek): a Greek god who shaped humans out of mud and gave them the ability to create fire; *archaeum* from *archaea* (Greek): an ancient life. The genus name is an analogy between the evolutionary relationship this organism and the origin of eukaryotes, and the involvement of Prometheus in the origin of humans from sediments and the acquisition of an unprecedented oxygen-driven energy-harnessing ability. The species name, *syntrophicum*, *syn* (Greek): together with; *trephein* (Greek) nourish; *icus* (Latin) pertaining to. The species name refers to the syntrophic substrate utilization property of this strain.

**Locality.** Isolated from deep-sea methane-seep sediment of the Nankai Trough at 2,533 m water depth, off the Kumano area, Japan.

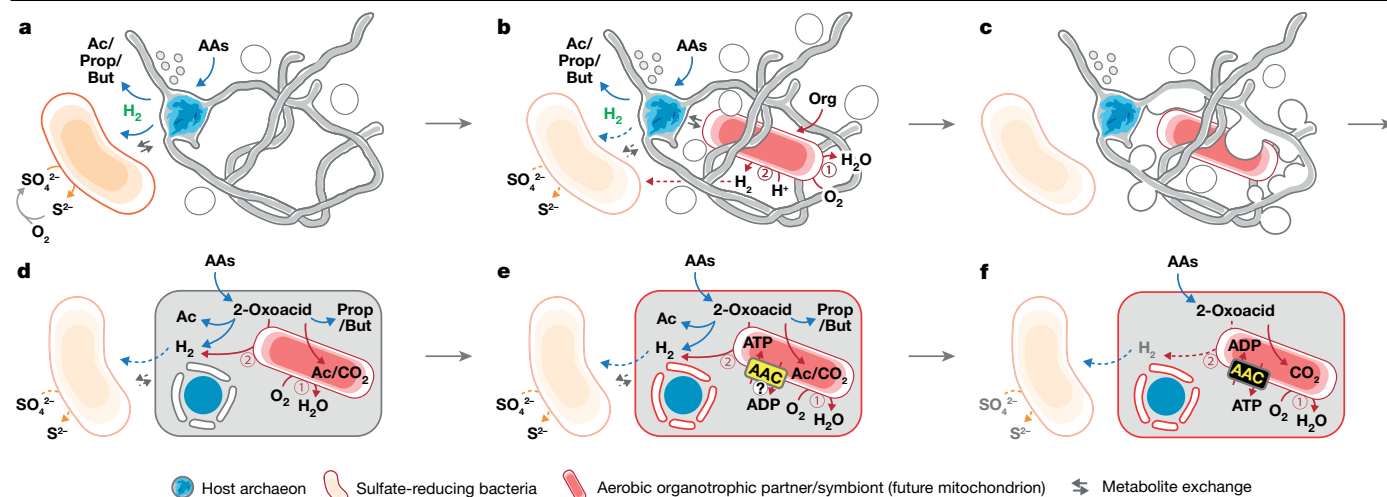
**Diagnosis.** Anaerobic, amino-acid-oxidizing archaeon, small coccus, around 550 nm in diameter, syntrophically grows with hydrogen- and formate-using microorganisms. It produces membrane vesicles, chains of blebs and membrane-based protrusions.

### Extant and ancestral features

The evolutionary relationship between archaea and eukaryotes has been under debate, hinging on the incompleteness and contamination associated with metagenome-derived genomes and variation in results that depend on tree-construction protocols<sup>22–25</sup>. By isolating MK-D1, we were able to obtain a closed genome (Extended Data Fig. 2 and Supplementary Table 2) and construct ribosomal protein-based phylogenomic trees that show clear a phylogenetic sister relation between MK-D1 and Eukarya (Fig. 4a, Extended Data Fig. 4 and Supplementary Tables 7, 8). Thus, MK-D1 represents the closest cultured archaeal relative of eukaryotes. We confirmed the presence of 80 eukaryotic signature proteins, which are also observed in related Asgard archaea (Supplementary Figs. 3–13 and Supplementary Tables 3, 9). Moreover, RNA-based evidence for expression of such genes was obtained. Among eukaryotic signature proteins, 23 fall in the 500 most highly expressed genes, including hypothetical proteins related to actin, gelsolin, ubiquitin, ESCRT-III proteins (Vps2/24/46-like and Vps20/32/60-like), Roadblock/







**Fig. 5 | Proposed hypothetical model for eukaryogenesis.** **a**, The syntrophic/fermentative host archaeon is suggested to degrade amino acids to short-chain fatty acids and  $H_2$ , possibly by interacting with  $H_2$ -scavenging (and indirectly  $O_2$ -scavenging) SRB (orange; see Supplementary Note 6). **b**, The host may have further interacted with a facultatively aerobic organotrophic partner that could scavenge toxic  $O_2$  (the future mitochondrion; red). Continued interaction with SRB could have been beneficial but not necessarily essential; dotted arrows indicate the interaction; see Supplementary Note 7. **c**, Host external structures could have interacted (for example, mechanical or

biological fusion<sup>50</sup>) with the aerobic partner to enhance physical interaction and further engulf the partner for simultaneous development of endosymbiosis and a primitive nucleoid-binding membrane. **d**, After engulfment, the host and symbiont could have continued the interaction shown in **b** as a primitive type of endosymbiosis. **e**, Development of ADP/ATP carrier (AAC) by the endosymbiont (initial direction of ATP transport remains unclear; see Supplementary Note 9). **f**, Endogenization of partner symbiosis by the host through delegation of catabolism and ATP generation to the endosymbiont and establishment of a symbiont-to-host ATP channel.

conserve related fermentative metabolic features across the superphylum, suggesting that the ancestor of the Asgard archaea possessed such capacities. This shows some congruence with a previous study that proposes hydrogenogenesis as a feature of the ancestor<sup>12</sup>, but differs in several central features.

## New insights into eukaryogenesis

The origin of the eukaryotic cell is one of the most enigmatic questions in biology. Isolation and cultivation of MK-D1 brings us closer to understanding how eukaryotes may have emerged; however, it is important to emphasize that the vast amount of time (roughly 2 billion years) that separates this modern-day organism from the organism that evolved into the last eukaryotic common ancestor (LECA) leaves many uncertainties—although we can make reasoned assumptions on the events that may have occurred during the course of evolution. The discussion that follows is a hypothetical model, in which we build on existing hypotheses with extrapolations from the insights gained in this study; notably, the model is not definitive and more studies on Asgard archaea and other deep-branching eukaryotes are required to contextualize the most probable steps that occurred.

Assuming that the ancestor of the Asgard archaea was indeed syntrophic, internally simple (that is, similar to MK-D1) and inhabited anaerobic marine sediments as most of the extant members of this lineage do<sup>6</sup>, evolution towards the facultatively aerobic LECA<sup>29</sup> can be envisioned to require (1) transition from anaerobiosis to aerobiosis, (2) the gain of an  $O_2$ -respiring and ATP-providing endosymbiont (that is, mitochondrion), and (3) development of intracellular structures. As Earth's  $O_2$  levels<sup>30</sup> had begun to rise before the evolution of the LECA (the TACK-Asgard archaea lineage dated to approximately 2.1–2.4 billion years ago<sup>31</sup>), we work on the assumption that the archaea needed to accommodate the increasing  $O_2$  levels, and energy and organic substrate inputs<sup>32</sup>, especially in benthic habitats of shallow oceans. Aero-tolerance might have been conferred by a symbiotic interaction with facultative  $O_2$ -respiring organisms<sup>33,34</sup>, which was potentially followed by endosymbiosis of one of these aerobes (that is, the future mitochondrion). Although such a transition from syntrophy to aerobiosis

is non-trivial, we suggest that a syntrophic interaction with SRB could have mediated this (Fig. 5a, b and Supplementary Notes 6, 7). Although previous models propose that  $H_2$  transfer was a key interaction that drove endosymbiosis<sup>12,29,35,36</sup>, we believe that current data favours the above interaction (see Supplementary Note 8). Given the small cell size of MK-D1 and the proposed lack of sufficient machinery<sup>37</sup> and energy<sup>38</sup>, we suggest that the physical manifestation of this endosymbiosis was probably independent of phagocytosis<sup>6</sup>. The observed morphology of strain MK-D1 rather points to a previously proposed alternative route<sup>39</sup> in which the host archaeon engulfed the metabolic partner using extracellular structures and simultaneously formed a primitive chromosome-surrounding structure that is topologically similar to the nuclear membrane; however, further evidence is required to support this conjecture (Fig. 5c, d).

After engulfment, the host may have shared amino-acid-derived 2-oxoacids with the endosymbiont as energy sources (Fig. 5d), given that amino-acid-degrading pathways widely encoded by Asgard archaea primarily recover ATP from 2-oxoacid degradation (Fig. 4b) and extant eukaryotes and mitochondria share 2-oxoacids<sup>40</sup>. In return, the endosymbiont may have consumed  $O_2$  (as proposed previously<sup>33</sup>) and provided the host with an intracellular pool of biological building blocks (for example, amino acids and co-factors that the host may not have been able to synthesize that were released passively or through endosymbiont death). On the basis of the absence of host-derived (that is, archaea-related) anaerobic 2-oxoacid catabolism genes (for example, ferredoxin-dependent 2-oxoacid oxidoreductase and NiFe hydrogenases) in eukaryotes<sup>41,42</sup>, the host presumably lost these during evolution towards the LECA. Notably, this loss might have consequently helped to simultaneously resolve catabolic redundancy (that is, 2-oxoacid catabolism in both host and symbiont) and  $O_2$  sensitivity (that is,  $O_2$  inactivates these enzymes<sup>43,44</sup>). For the resulting delegation of 2-oxoacid catabolism (and thus ATP generation) to the endosymbiont (as in modern mitochondria) to succeed, an ATP transport mechanism would have been necessary. Consistent with this notion, evolution of the ATP transporter (that is, the ADP/ATP carrier<sup>45</sup>) is thought to have been instrumental in fixing the symbiosis<sup>46</sup> (see Supplementary Note 9 for potential impetus; Fig. 5e). Another transition may have been

necessary—the host archaeon may have possessed ether-type lipids as observed for MK-D1 (Fig. 3j) and Asgard archaea<sup>47</sup>, yet all extant eukaryotes use ester-type lipids. However, a recent study showed that lipid types can mix without losing membrane integrity<sup>48</sup>, suggesting that the simple replacement of host ether-type lipids with ester-type lipids may have been possible (Fig. 5e). This hypothetical evolutionary scenario may have provided the steps that are required for the emergence of an aerobic organotroph that possess an O<sub>2</sub>-respiring ATP-generating endosymbiont congruent with extant eukaryotes and their mitochondria in terms of energy metabolism (Fig. 5f).

In summary, we have isolated and cultivated the closest archaeal relative of eukaryotes to date that has a unique metabolism and morphology, and combining these observations with genomic analyses, propose the entangle–engulf–endogenize model as one of several conceivable scenarios to explain the emergence of eukaryotes. Further investigation of MK-D1, related Asgard archaea and more deep-branching eukaryotes is now required and can provide valuable insights into the timing and progression of lateral gene transfer, endosymbiont organogenesis towards the first mitochondrion and the formation of the endomembrane system (among many other physiological features). Such endeavours are essential to refine our understanding of the possible chain of events that led to the eukaryotic cell, and to provide the necessary data that support or refute our models of eukaryogenesis.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-019-1916-6>.

- López-García, P. & Moreira, D. Open questions on the origin of eukaryotes. *Trends Ecol. Evol.* **30**, 697–708 (2015).
- Martin, W. F., Garg, S. & Zimorski, V. Endosymbiotic theories for eukaryote origin. *Phil. Trans. R. Soc. Lond. B* **370**, 20140330 (2015).
- Eme, L., Spang, A., Lombard, J., Stairs, C. W. & Ettema, T. J. G. Archaea and the origin of eukaryotes. *Nat. Rev. Microbiol.* **15**, 711–723 (2017).
- Koonin, E. V. Origin of eukaryotes from within archaea, archaeal eukaryome and bursts of gene gain: eukaryogenesis just made easier? *Phil. Trans. R. Soc. Lond. B* **370**, 20140333 (2015).
- Spang, A. et al. Complex archaea that bridge the gap between prokaryotes and eukaryotes. *Nature* **521**, 173–179 (2015).
- Zaremba-Niedzwiedzka, K. et al. Asgard archaea illuminate the origin of eukaryotic cellular complexity. *Nature* **541**, 353–358 (2017).
- Sousa, F. L., Neukirchen, S., Allen, J. F., Lane, N. & Martin, W. F. Lokiarchaeon is hydrogen dependent. *Nat. Microbiol.* **1**, 16034 (2016).
- Seitz, K. W., Lazar, C. S., Hinrichs, K.-U., Teske, A. P. & Baker, B. J. Genomic reconstruction of a novel, deeply branched sediment archaeal phylum with pathways for acetogenesis and sulfur reduction. *ISME J.* **10**, 1696–1705 (2016).
- Dombrowski, N., Teske, A. P. & Baker, B. J. Expansive microbial metabolic versatility and biodiversity in dynamic Guaymas Basin hydrothermal sediments. *Nat. Commun.* **9**, 4999 (2018).
- Liu, Y. et al. Comparative genomic inference suggests mixotrophic lifestyle for Thorarchaeota. *ISME J.* **12**, 1021–1031 (2018).
- Seitz, K. W. et al. Asgard archaea capable of anaerobic hydrocarbon cycling. *Nat. Commun.* **10**, 1822 (2019).
- Spang, A. et al. Proposal of the reverse flow model for the origin of the eukaryotic cell based on comparative analyses of Asgard archaeal metabolism. *Nat. Microbiol.* **4**, 1138–1148 (2019).
- Pushkarev, A. et al. A distinct abundant group of microbial rhodopsins discovered using functional metagenomics. *Nature* **558**, 595–599 (2018).
- Bulzu, P.-A. et al. Casting light on Asgardarchaeota metabolism in a sunlit microoxic niche. *Nat. Microbiol.* **4**, 1129–1137 (2019).
- Aoki, M. et al. A long-term cultivation of an anaerobic methane-oxidizing microbial community from deep-sea methane-seep sediment using a continuous-flow bioreactor. *PLoS ONE* **9**, e105356 (2014).
- Schink, B. & Stams, A. J. In *The Prokaryotes: Prokaryotic Communities and Ecophysiology* (eds Rosenberg, E. et al.) 471–493 (Springer, 2013).
- Knittel, K., Lösekann, T., Boetius, A., Kort, R. & Amann, R. Diversity and distribution of methanotrophic archaea at cold seeps. *Appl. Environ. Microbiol.* **71**, 467–479 (2005).
- Albers, S.-V. & Meyer, B. H. The archaeal cell envelope. *Nat. Rev. Microbiol.* **9**, 414–426 (2011).
- Marguet, E. et al. Membrane vesicles, nanobuds and/or nanotubes produced by hyperthermophilic archaea of the genus *Thermococcus*. *Biochem. Soc. Trans.* **41**, 436–442 (2013).
- Rosenshine, I., Tchelet, R. & Mevarech, M. The mechanism of DNA transfer in the mating system of an archaeobacterium. *Science* **245**, 1387–1389 (1989).
- Imachi, H. et al. Cultivation of methanogenic community from subseafloor sediments using a continuous-flow bioreactor. *ISME J.* **5**, 1913–1925 (2011).
- Da Cunha, V., Gaia, M., Gabelle, D., Nasir, A. & Forterre, P. Lokiarchaea are close relatives of Euryarchaeota, not bridging the gap between prokaryotes and eukaryotes. *PLoS Genet.* **13**, e1006810 (2017).
- Da Cunha, V., Gaia, M., Nasir, A. & Forterre, P. Asgard archaea do not close the debate about the universal tree of life topology. *PLoS Genet.* **14**, e1007215 (2018).
- Spang, A. et al. Asgard archaea are the closest prokaryotic relatives of eukaryotes. *PLoS Genet.* **14**, e1007080 (2018).
- Brunk, C. F. & Martin, W. F. Archaeal histone contributions to the origin of eukaryotes. *Trends Microbiol.* **27**, 703–714 (2019).
- Buckel, W. & Thauer, R. K. Energy conservation via electron bifurcating ferredoxin reduction and proton/Na<sup>+</sup> translocating ferredoxin oxidation. *Biochim. Biophys. Acta* **1827**, 94–113 (2013).
- Ma, K., Zhou, H. Z. & Adams, M. W. W. Hydrogen production from pyruvate by enzymes purified from the hyperthermophilic archaeon, *Pyrococcus furiosus*: a key role for NADPH. *FEMS Microbiol. Lett.* **122**, 245–250 (1994).
- Nobu, M. K. et al. The genome of *Syntrophorhabdus aromaticivorans* strain UI provides new insights for syntrophic aromatic compound metabolism and electron flow. *Environ. Microbiol.* **17**, 4861–4872 (2015).
- Martin, W. & Müller, M. The hydrogen hypothesis for the first eukaryote. *Nature* **392**, 37–41 (1998).
- Lyons, T. W., Reinhard, C. T. & Planavsky, N. J. The rise of oxygen in Earth's early ocean and atmosphere. *Nature* **506**, 307–315 (2014).
- Davin, A. A. et al. Gene transfers can date the tree of life. *Nat. Ecol. Evol.* **2**, 904–909 (2018).
- Kump, L. R. et al. Isotopic evidence for massive oxidation of organic matter following the great oxidation event. *Science* **334**, 1694–1696 (2011).
- Andersson, S. G. & Kurland, C. G. Origins of mitochondria and hydrogenosomes. *Curr. Opin. Microbiol.* **2**, 535–541 (1999).
- Fenchel, T. & Finlay, B. J. Oxygen toxicity, respiration and behavioural responses to oxygen in free-living anaerobic ciliates. *J. Gen. Microbiol.* **136**, 1953–1959 (1990).
- Moreira, D. & López-García, P. Symbiosis between methanogenic archaea and  $\delta$ -proteobacteria as the origin of eukaryotes: the syntrophic hypothesis. *J. Mol. Evol.* **47**, 517–530 (1998).
- López-García, P. & Moreira, D. Selective forces for the origin of the eukaryotic nucleus. *BioEssays* **28**, 525–533 (2006).
- Burns, J. A., Pittis, A. A. & Kim, E. Gene-based predictive models of trophic modes suggest Asgard archaea are not phagocytotic. *Nat. Ecol. Evol.* **2**, 697–704 (2018).
- Martin, W. F., Tielens, A. G. M., Mentel, M., Garg, S. G. & Gould, S. B. The physiology of phagocytosis in the context of mitochondrial origin. *Microbiol. Mol. Biol. Rev.* **81**, e00008-17 (2017).
- Baum, D. A. & Baum, B. An inside-out origin for the eukaryotic cell. *BMC Biol.* **12**, 76 (2014).
- Hutson, S. M. & Rannels, S. L. Characterization of a mitochondrial transport system for branched chain  $\alpha$ -keto acids. *J. Biol. Chem.* **260**, 14189–14193 (1985).
- Hug, L. A., Stechmann, A. & Roger, A. J. Phylogenetic distributions and histories of proteins involved in anaerobic pyruvate metabolism in eukaryotes. *Mol. Biol. Evol.* **27**, 311–324 (2010).
- Degli Esposti, M. et al. Alpha proteobacterial ancestry of the [Fe–Fe]-hydrogenases in anaerobic eukaryotes. *Biol. Direct* **11**, 34 (2016).
- Pieulle, L. et al. Isolation and characterization of the pyruvate-ferredoxin oxidoreductase from the sulfate-reducing bacterium *Desulfovibrio africanus*. *Biochim. Biophys. Acta* **1250**, 49–59 (1995).
- Liebgott, P.-P. et al. Relating diffusion along the substrate tunnel and oxygen sensitivity in hydrogenase. *Nat. Chem. Biol.* **6**, 63–70 (2010).
- Winkler, H. H. & Neuhaus, H. E. Non-mitochondrial ATP transport. *Trends Biochem. Sci.* **24**, 64–68 (1999).
- Gray, M. W. The pre-endosymbiont hypothesis: a new perspective on the origin and evolution of mitochondria. *Cold Spring Harb. Perspect. Biol.* **6**, a016097 (2014).
- Villanueva, L., Schouten, S. & Damsté, J. S. S. Phylogenomic analysis of lipid biosynthetic genes of Archaea shed light on the 'lipid divide'. *Environ. Microbiol.* **19**, 54–69 (2017).
- Caforio, A. et al. Converting *Escherichia coli* into an archaeobacterium with a hybrid heterochiral membrane. *Proc. Natl Acad. Sci. USA* **115**, 3704–3709 (2018).
- Nakamura, K. et al. Application of pseudomurein endoisopeptidase to fluorescence in situ hybridization of methanogens within the family *Methanobacteriaceae*. *Appl. Environ. Microbiol.* **72**, 6907–6913 (2006).
- Cevc, G. & Richardsen, H. Lipid vesicles and membrane fusion. *Adv. Drug Deliv. Rev.* **38**, 207–232 (1999).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020

# Article

## Methods

No statistical methods were used to predetermine sample size. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment.

### Sampling site and sample description

A 25-cm long sediment core (949C3) was collected from a methane-seep site at the Omine Ridge, Nankai Trough, off the Kumano area, Japan (33° 7.2253' N, 136° 28.6672' E), 2,533 m below sea level, by the manned submersible RV *Shinkai 6500* (cruise YK06-03, dive no. 6K949, 6 May 2006). The detailed sediment core sample and site information has been published previously<sup>15,51,52</sup>. Our previous geochemical and 16S rRNA gene analysis indicated that the occurrence of anaerobic oxidation of methane reactions was mediated by archaeal anaerobic methanotrophs in the sediment<sup>15,51</sup>. The SSU rRNA gene analysis also showed that the sediment contained abundant and diverse microorganisms, most of which were affiliated with uncultured microbial groups, including Asgard archaea<sup>15,51</sup>.

### Culturing

The deep-sea methane-seep sediment sample was first enriched using a continuous-flow bioreactor system supplemented with methane as the major energy source. The bioreactor, called a down-flow hanging sponge (DHS) bioreactor, has been operated in our laboratory, JAM-STE, Yokosuka Headquarters, since 28 December 2006. The detailed operation conditions for the DHS bioreactor have been described previously<sup>15</sup>. To isolate anaerobic microorganisms, including Asgard archaea, from the DHS reactor, 2-ml samples of the bioreactor enrichment sediment slurry were inoculated in 15-ml glass tubes with a simple substrate and a basal medium. The composition of the basal medium was almost similar to that used for cultivation in the DHS bioreactor<sup>15</sup>, but it did not contain sulfate (that is, Na<sub>2</sub>SO<sub>4</sub>). The basal medium composition was as follows (per litre): 9.47 g MgCl<sub>2</sub>·6H<sub>2</sub>O, 1.36 g CaCl<sub>2</sub>·2H<sub>2</sub>O, 20.7 g NaCl, 0.54 g NH<sub>4</sub>Cl, 0.14 g KH<sub>2</sub>PO<sub>4</sub>, 2.7 g NaHCO<sub>3</sub>, 0.3 g Na<sub>2</sub>S·9H<sub>2</sub>O, 0.3 g cysteine-HCl, 1 ml trace element solution<sup>15</sup>, 1 ml Se/W solution, 2 ml vitamin solution<sup>15</sup> and resazurin solution (1 mg ml<sup>-1</sup>). The medium was purged with N<sub>2</sub>/CO<sub>2</sub> gas (80:20, v/v), and the pH was adjusted to 7.5 at 25 °C. The culture tubes were sealed with butyl rubber stoppers and screw caps. Autoclaved or filter-sterilized organic substances (such as protein-derived materials, sugars and fatty acids) were added to the tubes with stock solutions before inoculation with the bioreactor-enriched community. After establishing a stable *Ca. P. syntrophicum* culture, cultivations were performed at 20 °C in 50-ml serum vials containing 20 ml basal medium supplemented with casamino acids (0.05%, w/v), 20 amino acids (0.1 mM each) and powdered milk (0.1%, w/v, Hohoemi, Meiji) under an atmosphere of N<sub>2</sub>/CO<sub>2</sub> (80:20, v/v) in the dark without shaking, unless mentioned otherwise. Information regarding the purity check of MK-D1 cultures, as well as additional information about cultivation, is included in the Supplementary Methods.

### SSU rRNA gene-based analysis

DNA extraction and PCR mixture preparation were performed on a clean bench to reduce contamination. DNA extraction from culture samples was performed as described previously<sup>53</sup>. The concentration of extracted DNA was measured using a Quant-iT dsDNA High-Sensitivity Assay Kit (Life Technologies). PCR amplification was performed using the Takara Ex *Taq* (for conventional clone analysis) or Takara LA *Taq* (for Illumina-based amplicon sequencing (iTAG) for targeted sequencing for the SSU rRNA gene analysis) (Takara Bio), and the reaction mixtures for PCR were prepared according to the manufacturer's instructions. For the conventional clone analysis, a universal primer pair 530F/907R<sup>51</sup> and an archaeal primer pair 340F/932R<sup>15,54</sup> were used for PCR amplification. For iTAG analysis, the universal primer pair 530F/907R, which contained overhang adapters at the 5' ends, was used. The procedures used for

library construction, sequencing and data analysis were described previously<sup>21,55</sup>.

### Growth monitoring using qPCR

For the quantitative analysis, a StepOnePlus Real-Time PCR System (Thermo Fisher Scientific) with a SYBR Premix Ex Taq II kit (TaKaRa Bio) was used. The candidate phylum Lokiarchaeota-specific primer pair MBGB525F/Ar912r was used for amplification of 16S rRNA genes. Primer MBGB525F is the complementary sequence of the MGBG525 probe<sup>17</sup>, whereas Ar912r is an archaeal universal primer that is a slightly modified version of the originally designed primer<sup>56</sup>. The detailed procedure for qPCR is described in the Supplementary Methods. The doubling times of MK-D1 were calculated based on the semi-logarithmic plot of the qPCR data.

### Growth test with multiple substrates

To examine the effect of the presence of other substances on the growth of MK-D1, medium containing casamino acids, 20 amino acids, powdered milk and supplemented with an individual substrate (Extended Data Table 3) was prepared, followed by qPCR and iTAG analyses. Each cultivation condition was set in duplicate; however, the H<sub>2</sub>-fed culture was prepared in triplicate because a previous study<sup>7</sup> reported that a Lokiarchaeum has potential to grow with hydrogen based on a comparative genome analysis. Detailed culture liquid sampling and the subsequent qPCR and iTAG analyses are described in the Supplementary Information.

### Evaluation of growth temperature

The test was performed using a basal medium containing casamino acids and powdered milk, with a pure co-culture of MK-D1 and *Methanogenium* as the inoculum (20%, v/v). The cultures were incubated at 4, 10, 15, 20, 25, 30, 37 and 40 °C. All incubations for the test were performed in triplicate. After 100 days of incubation, 16S rRNA gene copy numbers of MK-D1 were evaluated using qPCR.

### FISH

Fixation of microbial cells, storage of the fixed cells and standard FISH were performed in accordance with a previously described protocol<sup>21</sup>. The 16S rRNA-targeted oligonucleotide probes used in this study are listed in Supplementary Table 10. The design of MK-D1-specific probes is described in the Supplementary Methods. As clear fluorescent signals were not obtained using the standard FISH technique, we used an in situ DNA-hybridization chain reaction (HCR) technique<sup>57</sup>. The FISH samples were observed using epifluorescence microscopes (BX51 or BX53, Olympus) and a confocal laser scanning microscope (Nikon A1RMP, Nikon Instech).

### SEM

Microbial cells were fixed overnight in 2.5% (w/v) glutaraldehyde in the casamino acids–20 amino acid medium at 20 °C. The sample preparation procedure has been described previously<sup>58</sup>. The cell samples were observed using field emission-SEM (JSM-6700F, JEOL) or extreme high-resolution FIB-SEM (Helios G4 UX, ThermoFisher Scientific).

### Ultrathin sectioning and TEM

Cells were prefixed with 2.5% (w/v) glutaraldehyde for 2 h. The specimens were frozen in a high-pressure freezing apparatus (EM-PACT2, Leica)<sup>59</sup>. The frozen samples were substituted with 2% OsO<sub>4</sub> in acetone for 3–4 days at –80 °C, and the samples were warmed gradually to room temperature, rinsed with acetone embedded in epoxy resin (TAAB). Thin sections (70 nm) were cut with an ultramicrotome (EM-UC7, Leica). Ultrathin sections of the cells were stained with 2% uranyl acetate and lead-stained solution (0.3% lead nitrate and 0.3% lead acetate, Sigma-Aldrich), and were observed using TEM (Tecnai 20, FEI) at an acceleration voltage of 120 kV.



## Cryo-EM

Owing to the low cell yield culture, 400 ml of the culture of MK-D1 was prepared and concentrated to about 5 ml using a 0.22- $\mu$ m-pore-size polyethersulfone filter unit (Corning) in an anaerobic chamber (95:5 (v/v) N<sub>2</sub>:H<sub>2</sub> atmosphere; COY Laboratory Products). The concentrated culture liquid was placed in a glass vial in the anaerobic chamber. After that, the head space of the glass vial was replaced by N<sub>2</sub>/CO<sub>2</sub> gas (80:20, v/v). Immediately before the observation using electron microscopy, the glass vial was opened, and the liquid culture was concentrated to about 200  $\mu$ l by centrifugation at 20,400g for 10 min at 20 °C. Subsequently, 3  $\mu$ l of the concentrated liquid culture was applied onto a Quantifoil Mo grid R1.2/1.3 (Quantifoil MicroTools) pretreated with glow-discharge, and was plunged-frozen in liquid ethane using a Vitrobot Mark IV (FEI Company) at 4 °C and 95% humidity.

The frozen grid was mounted onto a 914 liquid-nitrogen cryo-specimen holder (Gatan) and loaded into a JEM2200FS electron microscope (JEOL) equipped with a field emission electron source operating at 200 kV and an omega-type in-column energy filter (slit width: 20 eV). The images were recorded on a DE-20 direct detector camera (Direct Electron) at a nominal magnification of 15,000 $\times$ , which resulted in an imaging resolution of 3.66 Å per pixel, with the total dose under 20 electrons per Å<sup>2</sup> using a low-dose system. For electron tomography, tilt series images were collected manually in a range of approximately  $\pm 62^\circ$  at  $2^\circ$  increments. The total electron dose on the specimen per tilt series was kept under 100 electrons per Å<sup>2</sup> to minimize radiation damage. The tilt series were aligned using gold fiducials and tomograms were reconstructed using filtered back projection or SIRT in the IMOD software<sup>60</sup> with an image binning of 5.

## Lipid analysis

About 120 ml of a highly purified culture sample was concentrated using the same method as described above, except that the filtration concentration procedure was performed on a clean bench instead of the anaerobic chamber. After cell collection, the cells were washed with the anaerobic basal medium to eliminate the interfering matrix. Subsequently, lipid analysis was conducted for the collected cells after the improved method<sup>61</sup>. For precise qualitative liquid analysis, GC–MS was conducted on the 7890 system (Agilent Technologies) to compare the retention time and mass fragmentation signatures.

## Stable isotope probing and NanoSIMS analysis

To confirm utilization of amino acids by MK-D1, a stable-isotope probing experiment was performed using a <sup>13</sup>C- and <sup>15</sup>N-labelled amino acid mixture (Cambridge Isotope Laboratories). In brief, 120 ml serum vials containing 40 ml basal medium were prepared and supplemented with the 20 stable-isotope-labelled amino acids (roughly 0.1 mM of each), casamino acids (0.05%, w/v) and non-labelled 20 amino acid mixture (0.1 mM of each). Two types of highly purified cultures of MK-D1 were used as inocula: a co-culture with *Methanobacterium* sp. strain MO-MB1 and a tri-culture with *Halodesulfobivibrio* and *Methanogenium*. The vials were incubated at 20 °C in the dark without shaking for 120 days. A reference cultivation was also performed under the same cultivation conditions without the addition of the 20 stable-isotope-labelled amino acid mixture (Extended Data Table 2). The detailed sample preparation and analysis method using NanoSIMS is described in the Supplementary Methods.

## Chemical analysis

The stable carbon isotope compositions of methane and CO<sub>2</sub> in the sampled gas phase were analysed as described previously<sup>62</sup>. Methane concentrations were measured by GC (GC-4000, GL Science) using a Shincarbon ST 50/80 column (1.0 m  $\times$  3.0 mm inner diameter; Shinwa Chemical Industries) and a flame ionization detector with nitrogen as a carrier gas.

Amino acid concentrations in pure co-cultures of MK-D1 and *Methanogenium* were quantified through a previously described method<sup>63,64</sup>. In brief, we processed the acid hydrolysis with 6 M HCl (110 °C, 12 h) for the culture liquid samples after filtration using a 0.2- $\mu$ m pore-size polytetrafluoroethylene filter unit (Millipore). The amino acid fraction was derivatized to *N*-pivaloyl iso-propyl esters before GC using a 6890NGC instrument connected to the nitrogen phosphorus and flame ionization detectors (Agilent Technologies). For cross-validation of qualitative identification of amino acids, GC–MS on the 7890 system (Agilent Technologies) was used<sup>61</sup>.

## Genome sequencing and assembly

DNA extraction was performed as described previously<sup>53</sup>. Mate-paired library with an average insert size of 3,000 bp was constructed according to the manufacturer's instructions with Nextera Mate Pair Library Preparation kit (Illumina). Library sequencing was performed using Illumina MiSeq platform (2  $\times$  300 bp), which resulted in 3,822,290 paired reads. The mate pair reads were processed as follows: adapters and low-quality sequences were removed using Trimmomatic v.0.33<sup>63</sup> (ILLUMINACLIP:TruSeq3-PE-2.fa:2:30:10:8:true LEADING:3 TRAILING:3 SLIDINGWINDOW:4:20 MINLEN:100), and the linker sequences were removed using NextClip v.1.3.1<sup>65</sup>. De novo assembly was performed using SPAdes v.3.11<sup>66</sup> with multiple *k*-mer sizes (21, 33, 55, 77 and 99), which resulted in 3,487 contigs with lengths >500 bp, totalling up to 14.68 Mb. The software MyCC<sup>67</sup> was used with default parameters for binning based on genomic signatures, marker genes and contig coverages. As heterogeneity in the sequence can cause highly fragmented or redundant contigs, the ambiguous contigs (sequence coverage <5 or a length <1 kb) and redundant contigs were discarded from binning. This resulted in the recovery of genomes related to *Lokiarchaeota* (that is, *Ca. P. syntrophicum* MK-D1, 4.46 Mb), *Halodesulfobivibrio* (4.13 Mb) and *Methanogenium* (2.33 Mb). Scaffolds for each bin were constructed using SSPACE v.3.0<sup>68</sup> with mate-paired information of Illumina reads. To obtain the complete genome sequence of *Ca. P. syntrophicum*, the gaps were filled using Sanger sequencing. Genomes were annotated using Prokka v.1.12<sup>69</sup> and manually curated. The curation involved functional domain analysis through CD-Search (CDD v.3.17) with its corresponding conserved domain database<sup>70,71</sup> and InterProScan v.5<sup>72</sup>; signal peptide and transmembrane domain prediction through SignalP v.4.1<sup>73</sup>; carbohydrate-active enzyme, peptidase and lipase prediction through dbCAN v.5.0<sup>74</sup>, MEROPS<sup>75</sup> and lipase engineering database<sup>76</sup>; and hydrogenase annotation with assistance from HydDB<sup>77</sup>. In addition, to further verify the function, we compared the sequence similarity of each gene to enzymes found in UniProtKB/SwissProt that had experimentally verified catalytic activity and genes with extensive genetic, phylogenetic and/or genomic characterizations<sup>78,79</sup> with a 40% amino acid similarity cut-off. For enzymes that have divergent functions even with a 40% similarity cut-off (for example, [FeFe] and [NiFe] hydrogenases, 3-oxoacid oxidoreductases, glutamate dehydrogenases and sugar kinases), phylogenetic trees were constructed with reference sequences to identify association of the query sequences to phylogenetic clusters containing enzymes with characterized catalytic activity. Publicly available metagenome-assembled genomes of Asgard archaea were annotated in the same manner.

## Phylogenetic analysis

Phylogenomic trees of MK-D1 and select cultured archaea, eukaryotes and bacteria were calculated. Thirty-one ribosomal proteins conserved across the three domains (Supplementary Table 7) were collected from MK-D1, the organisms shown in the tree and metagenome-assembled genomes (MAGs) of uncultured archaeal lineages (Supplementary Table 8). Two alignments were performed in parallel: (1) only including sequences from cultured organisms and (2) also including MAG-derived sequences. MAFFT v.7 (--linsi) was used for alignment in both cases<sup>80</sup>. For the latter, MAG-derived sequences were included to generate an



# Article

alignment that maximizes the archaeal diversity that is taken into account, but removed for subsequent tree construction to avoid any influence of contamination (that is, concatenation of sequences that do not belong to the same organism). ‘*Candidatus* Korarchaeum’ sequences were kept in the tree based on the cultured + uncultured alignment due to its critical position in TACK phylogeny. After removing all-gap positions and concatenation, the maximum-likelihood trees were constructed using RAXML-NG v.0.8.0<sup>81</sup> (fixed empirical substitution matrix (LG), 4 discrete GAMMA categories, empirical amino acid frequencies and 100 bootstrap replicates) and the Bayesian inference phylogenies were calculated using MrBayes v.3.2.7a<sup>82</sup> (four chains, print/sample frequencies of 100, a relative burn-in of 25% (nchains = 4 nruns = 2 printfreq = 100 samplefreq = 100), LG model, invariable sites plus GAMMA models of rate variation across sites (prset aamodelpr = fixed(lg); lset rates = invgamma)). For 16S ribosomal RNA phylogeny, sequences were aligned using SINA<sup>83</sup> against the Silva v.132 alignment<sup>84</sup>. The maximum-likelihood tree was calculated using RAXML<sup>85</sup> using the same parameters as RAXML-NG.

For analysis of urocanate hydratase, serine/threonine dehydratase, succinate dehydrogenase flavoprotein, fatty-acid-CoA ligase and 3-ketoacyl-CoA thiolase, homologues were collected through BLASTp<sup>86</sup> analysis of the Asgard archaea sequences against the UniProt database (release 2019\_05). Asgard archaea protein sequences unavailable in GenBank or UniProt (that is, those without accession numbers in the trees) were predicted with Prokka v.1.13<sup>69</sup> (--kingdom Archaea --rnammer) using the genome assemblies available in GenBank. Of homologues with sequence similarity  $\geq 40\%$  and overlap  $\geq 70\%$ , representative sequences were selected using CD-HIT v.4.8.1<sup>87</sup> with a clustering cut-off of 70% similarity (default settings otherwise). Additional homologues with verified biochemical activity, sequence similarity  $\geq 30\%$ , and overlap  $\geq 70\%$  were collected through BLASTp<sup>86</sup> analysis of the Asgard archaea sequences against the UniProt/SwissProt database (2019\_05)<sup>88</sup>. Sequences were aligned using MAFFT v.7<sup>80</sup> with default settings (or MUSCLE v.3.8.31<sup>89</sup> where noted) and trimmed using trimAl v.1.2<sup>90</sup> (settings are specified in the caption for each corresponding phylogenetic tree). RAXML-NG<sup>81</sup> was used for tree construction with the same parameters above (or PhyML v.3.3<sup>91</sup> with 100 bootstrap replicates, LG model and empirical amino acid frequencies where noted). For analysis of biotin ligase and biotin carboxyl carrier protein, the phylogenetic tree was constructed using FastTree<sup>92</sup> using the LG model and 1,000 bootstrap replicates.

## RNA-based sequencing analysis

To perform RNA-based sequencing analysis, 100 ml of culture liquid was prepared from 5 highly purified cultures that were incubated with casamino acids, 20 amino acids and powdered milk for about 100 days at 20 °C. Before RNA extraction, the growth of MK-D1 was confirmed using qPCR, and the cells density levels were around  $10^5$  copies ml<sup>-1</sup> in each culture.

To collect microbial cells, the culture liquid was filtered through a 0.22- $\mu$ m pore-size mixed cellulose ester membrane filter (GSPW01300, Merck MilliPore) on a clean bench. After filtration, the membrane was cut in half with sterilized scissors and then directly inserted into the PowerBiofilm bead tubes of a PowerBiofilm RNA Isolation kit (MO BIO Laboratories). The following RNA extraction procedures were performed according to the manufacturer’s instructions. The extracted RNA was applied to an RNA Clean & Concentrator Kit-5 (Zymo Research) for concentration. The obtained RNA was quantified using an Agilent 2100 Bioanalyzer system with an RNA Pico kit (Agilent Technologies) and then applied to an Ovation Universal RNA-Seq System (NuGEN Technologies) for the construction of an RNA-sequence library. At the step for Insert Dependent Adaptor Cleavage technology-mediated adaptor cleavage during the library construction, specific primers for 16S rRNA and 23S rRNA genes of MK-D1 were used to reduce rRNA gene sequences from the cDNA

pool. The constructed cDNA library was sequenced using the MiSeq platform (Illumina).

The raw RNA sequencing data were trimmed by removal of the adaptors and low-quality sequences using Trimmomatic v.0.33<sup>93</sup>. The expression abundance of all coding transcripts was estimated in RPKM values using EDGE-pro v.1.3.1<sup>94</sup>.

## Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

## Data availability

Genomes for *Ca. P. syntrophicum* MK-D1, *Halodesulfovibrio* sp. MK-HDV and *Methanogenium* sp. MK-MG are available under GenBank BioProject accession numbers PRJNA557562, PRJNA557563 and PRJNA557565, respectively. The iTAG sequence data was deposited in BioProject PRJDB8518 with SRA accession numbers DRR184081–DRR184101. The 16S rRNA gene sequences of MK-D1, *Halodesulfovibrio* sp. MK-HDV, *Methanogenium* sp. MK-MG and clones obtained from primary enrichment culture were deposited in the DDBJ/EMBL/GenBank database under accession numbers LC490619–LC490624. The gene expression data of MK-D1 in BioProject PRJDB9032 with the accession number DRR199588. The cryo-electron tomograms of *Ca. P. syntrophicum* MK-D1 have been deposited in the EMDB with accession codes EMD-0809 and EMD-0852.

51. Nunoura, T. et al. Microbial diversity in deep-sea methane seep sediments presented by SSU rRNA gene tag sequencing. *Microbes Environ.* **27**, 382–390 (2012).
52. Toki, T., Higa, R., Ijiri, A., Tsunogai, U. & Ashi, J. Origin and transport of pore fluids in the Nankai accretionary prism inferred from chemical and isotopic compositions of pore water at cold seep sites off Kumano. *Earth Planets Space* **66**, 137 (2014).
53. Nakahara, N. et al. *Aggregatilinea lenta* gen. nov., sp. nov., a slow-growing, facultatively anaerobic bacterium isolated from subsurface sediment, and proposal of the new order *Aggregatilineales* ord. nov. within the class *Anaerolineae* of the phylum *Chloroflexi*. *Int. J. Syst. Evol. Microbiol.* **69**, 1185–1194 (2019).
54. Murakami, S., Fujishima, K., Tomita, M. & Kanai, A. Metatranscriptomic analysis of microbes in an oceanfront deep-sea hydrothermal vent reveals novel small RNAs and type-specific tRNA degradation. *Appl. Environ. Microbiol.* **78**, 1015–1022 (2012).
55. Imachi, H. et al. Cultivable microbial community in 2-km-deep, 20-million-year-old subsurface coalbeds through -1000 days anaerobic bioreactor cultivation. *Sci. Rep.* **9**, 2305 (2019).
56. Miyashita, A. et al. Development of 16S rRNA gene-targeted primers for detection of archaeal anaerobic methanotrophs (ANMEs). *FEMS Microbiol. Lett.* **297**, 31–37 (2009).
57. Yamaguchi, T. et al. In situ DNA-hybridization chain reaction (HCR): a facilitated in situ HCR system for the detection of environmental microorganisms. *Environ. Microbiol.* **17**, 2532–2541 (2015).
58. Miyazaki, M. et al. *Sphaerochaeta multiformis* sp. nov., an anaerobic, psychrophilic bacterium isolated from subsurface sediment, and emended description of the genus *Sphaerochaeta*. *Int. J. Syst. Evol. Microbiol.* **64**, 4147–4154 (2014).
59. Toyooka, K. et al. Wide-range high-resolution transmission electron microscopy reveals morphological and distributional changes of endomembrane compartments during log to stationary transition of growth phase in tobacco BY-2 cells. *Plant Cell Physiol.* **55**, 1544–1555 (2014).
60. Kremer, J. R., Mastroratte, D. N. & McIntosh, J. R. Computer visualization of three-dimensional image data using IMOD. *J. Struct. Biol.* **116**, 71–76 (1996).
61. Takano, Y. et al. Insight into anaerobic methanotrophy from <sup>13</sup>C/<sup>12</sup>C- amino acids and <sup>14</sup>C/<sup>12</sup>C-ANME cells in seafloor microbial ecology. *Sci. Rep.* **8**, 14070 (2018).
62. Okumura, T. et al. Hydrogen and carbon isotope systematics in hydrogenotrophic methanogenesis under H<sub>2</sub>-limited and H<sub>2</sub>-enriched conditions: implications for the origin of methane and its isotopic diagnosis. *Prog. Earth Planet. Sci.* **3**, 14 (2016).
63. Takano, Y., Kashiwaga, Y., Ogawa, N. O., Chikaraishi, Y. & Ohkouchi, N. Isolation and desalting with cation-exchange chromatography for compound-specific nitrogen isotope analysis of amino acids: application to biogeochemical samples. *Rapid Commun. Mass Spectrom.* **24**, 2317–2323 (2010).
64. Chikaraishi, Y. et al. *Instrumental Optimization for Compound-specific Nitrogen Isotope Analysis of Amino Acids by Gas Chromatography/Combustion/Isotope Ratio Mass Spectrometry in Earth, Life and Isotopes* (eds Ohkouchi, N. et al.) 367–386 (Kyoto Univ. Press, 2010).
65. Leggett, R. M., Clavijo, B. J., Clissold, L., Clark, M. D. & Caccamo, M. NextClip: an analysis and read preparation tool for Nextera long mate pair libraries. *Bioinformatics* **30**, 566–568 (2014).
66. Bankevich, A. et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **19**, 455–477 (2012).
67. Lin, H.-H. & Liao, Y.-C. Accurate binning of metagenomic contigs via automated clustering sequences using information of genomic signatures and marker genes. *Sci. Rep.* **6**, 24175 (2016).

68. Boetzer, M., Henkel, C. V., Jansen, H. J., Butler, D. & Pirovano, W. Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* **27**, 578–579 (2011).
69. Seemann, T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068–2069 (2014).
70. Marchler-Bauer, A. & Bryant, S. H. CD-Search: protein domain annotations on the fly. *Nucleic Acids Res.* **32**, W327–W331 (2004).
71. Marchler-Bauer, A. et al. CDD: NCBI's conserved domain database. *Nucleic Acids Res.* **43**, D222–D226 (2015).
72. Jones, P. et al. InterProScan 5: genome-scale protein function classification. *Bioinformatics* **30**, 1236–1240 (2014).
73. Petersen, T. N., Brunak, S., von Heijne, G. & Nielsen, H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* **8**, 785–786 (2011).
74. Yin, Y. et al. dbCAN: a web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* **40**, W445–W451 (2012).
75. Rawlings, N. D., Barrett, A. J. & Finn, R. Twenty years of the MEROPS database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res.* **44**, D343–D350 (2016).
76. Fischer, M. & Pleiss, J. The Lipase Engineering Database: a navigation and analysis tool for protein families. *Nucleic Acids Res.* **31**, 319–321 (2003).
77. Søndergaard, D., Pedersen, C. N. S. & Greening, C. HydDB: a web tool for hydrogenase classification and analysis. *Sci. Rep.* **6**, 34212 (2016).
78. Boutet, E., Lieberherr, D., Tognolli, M., Schneider, M. & Bairoch, A. UniProtKB/Swiss-Prot. *Methods Mol. Biol.* **406**, 89–112 (2007).
79. Lima, T. et al. HAMAP: a database of completely sequenced microbial proteome sets and manually curated microbial protein families in UniProtKB/Swiss-Prot. *Nucleic Acids Res.* **37**, D471–D478 (2009).
80. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
81. Kozlov, A. M., Darriba, D., Flouri, T., Morel, B. & Stamatakis, A. RAXML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* **35**, 4453–4455 (2019).
82. Ronquist, F. et al. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **61**, 539–542 (2012).
83. Pruesse, E., Peplies, J. & Glöckner, F. O. SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* **28**, 1823–1829 (2012).
84. Quast, C. et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596 (2013).
85. Stamatakis, A. RAXML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).
86. Camacho, C. et al. BLAST+: architecture and applications. *BMC Bioinformatics* **10**, 421 (2009).
87. Fu, L., Niu, B., Zhu, Z., Wu, S. & Li, W. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* **28**, 3150–3152 (2012).
88. UniProt Consortium. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.* **47**, D506–D515 (2019).
89. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).
90. Capella-Gutiérrez, S., Silla-Martínez, J. M. & Gabaldón, T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972–1973 (2009).
91. Guindon, S. et al. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* **59**, 307–321 (2010).
92. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS ONE* **5**, e9490 (2010).
93. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
94. Magoc, T., Wood, D. & Salzberg, S. L. EDGE-pro: estimated degree of gene expression in prokaryotic genomes. *Evol. Bioinform. Online* **9**, 127–136 (2013).
95. Axley, M. J. & Grahame, D. A. Kinetics for formate dehydrogenase of *Escherichia coli* formate-hydrogenlyase. *J. Biol. Chem.* **266**, 13731–13736 (1991).
96. Itoh, T., Suzuki, K. & Nakase, T. *Thermocladium modestius* gen. nov., sp. nov., a new genus of rod-shaped, extremely thermophilic crenarchaeote. *Int. J. Syst. Bacteriol.* **48**, 879–887 (1998).
97. Zillig, W. et al. The archaeobacterium *Thermofilum pendens* represents, a novel genus of the thermophilic, anaerobic sulfur respiring *Thermoproteales*. *Syst. Appl. Microbiol.* **4**, 79–87 (1983).

**Acknowledgements** We thank H. Ohno and T. Yamaguchi for assistance with HCR-FISH analysis; T. Terada for help with NanoSIMS sample preparation; M. Isozaki for assistance with cultivation experiments; T. Kubota for assistance with chemical analysis; K. Takishita, A. Yabuki, T. Shiratori, A. Ohashi, F. Inagaki, T. Nunoura, S. Kawagucci, T. Shibuya, S. Ishii, S. Suzuki, Y. Tsukatani, C. Chen, Y. Kuruma and R. C. Robinson for advice and discussion; A. Miyashita, Y. Yashiro, K. Aoi, M. Ehara, M. Aoki and Y. Saito for assistance with operating the bioreactor; and J. Ashi and the RV *Yokosuka* and RV *Shinkai 6500* operation team during cruise YK06-03 (JAMSTEC) and the shipboard scientists and crews of the RV *Chikyu* Shakedown Cruise CK06-06 for their assistance in collecting samples. This study was partially supported by grants from the Japan Society for the Promotion of Science (JSPS) (KAKENHI grants 18687006, 21687006, 24687011, 15H02419 and 19H01005 to H.I., 18H03367 to M.K.N., 26710012, 18H02426, 18H05295 to H.T., 18H04468 and 18K18795 to M.I. and Grant-in-Aid for JSPS Fellow 16J10845 to N.N.). This work was also supported by JSPS KAKENHI grant number JP16H06280, Grant-in-Aid for Scientific Research on Innovative Areas—Platforms for Advanced Technologies and Research Resources 'Advanced Bioimaging Support' and the Cooperative Study Program (19-504) of National Institute for Physiological Sciences.

**Author contributions** H.I. conceived the study and carried out the deep-marine sediment sampling. H.I., N.N., M.O., M.M. and S.S. conducted cultivation and culture-based experiments. M.K.N. performed metabolic reconstruction and phylogenetic analyses. M.K.N. and Y. Takaki performed genome analysis. H.I., N.N., Y. Morono, M.O., T.I., M.I., K.M., C.S. and K.U. carried out the microscopy and NanoSIMS work. M.O., Y.S. and Y.Y. performed qPCR, SSU rRNA gene analysis and DNA/RNA sequencing. Y. Takano, Y. Matsui and E.T. performed chemical analysis. H.I., M.K.N., N.N., Y. Morono, Y. Takaki, Y. Takano, K.M., C.S., T.Y., Y.K., H.T. and K.T. conducted data interpretation. H.I., M.K.N., Y. Takano, H.T., Y.K. and K.T. wrote the manuscript with input from all co-authors. All authors have read and approved the manuscript submission.

**Competing interests** The authors declare no competing interests.

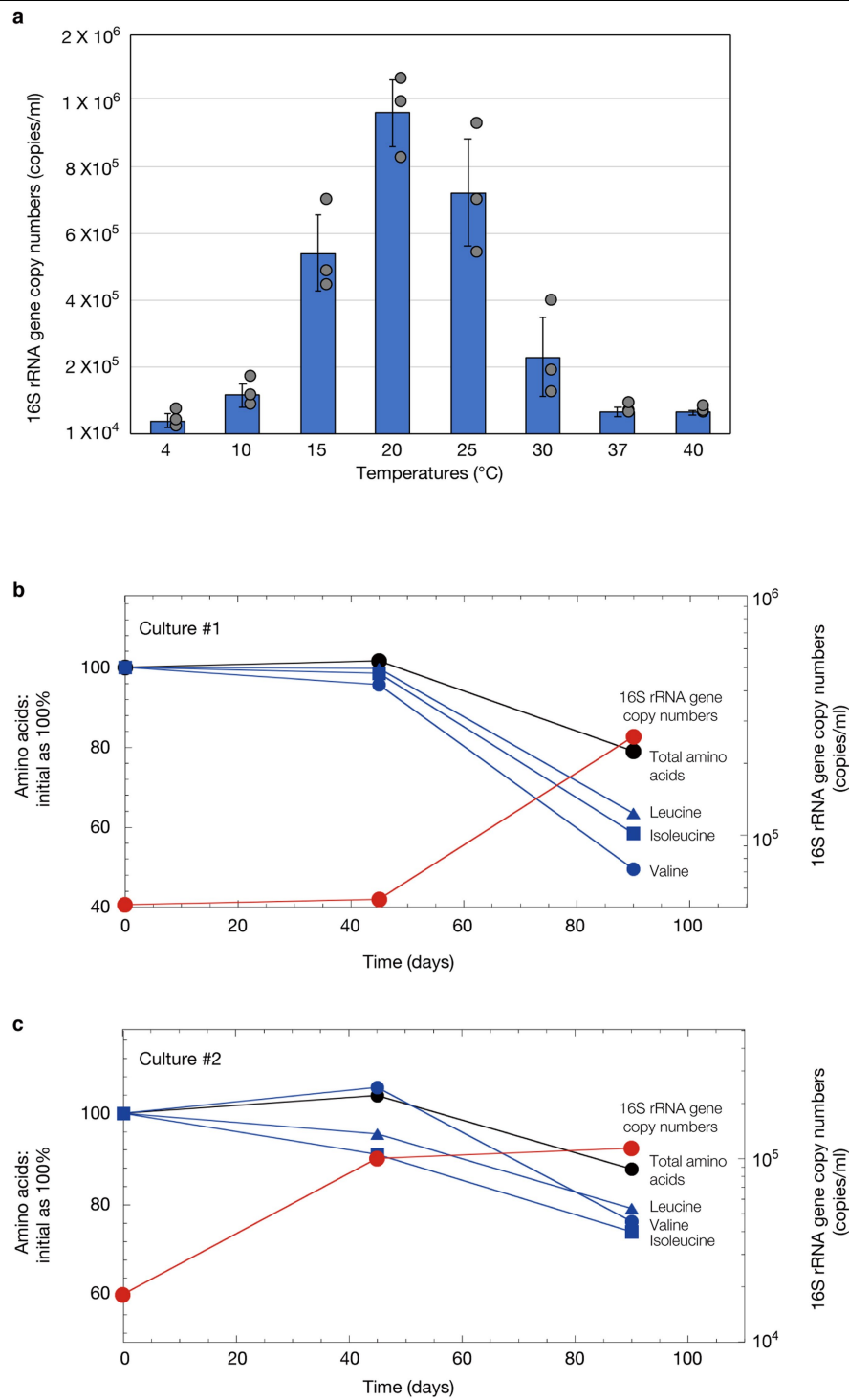
#### Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41586-019-1916-6>.

**Correspondence and requests for materials** should be addressed to H.I. or M.K.N.

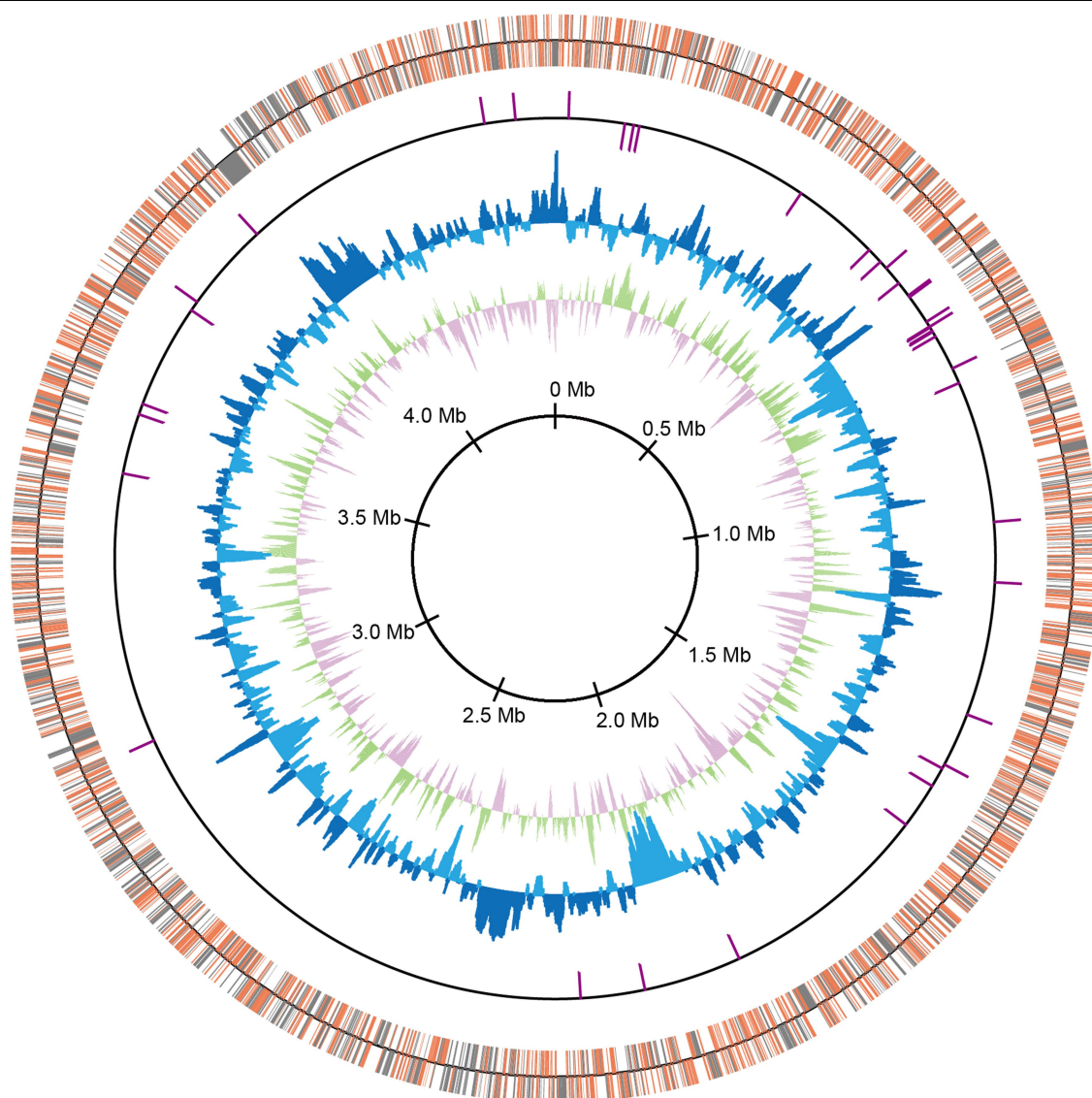
**Peer review information** *Nature* thanks Sonja-Verena Albers, Petr G. Leiman, James McInerney, Christa Schleper and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>.



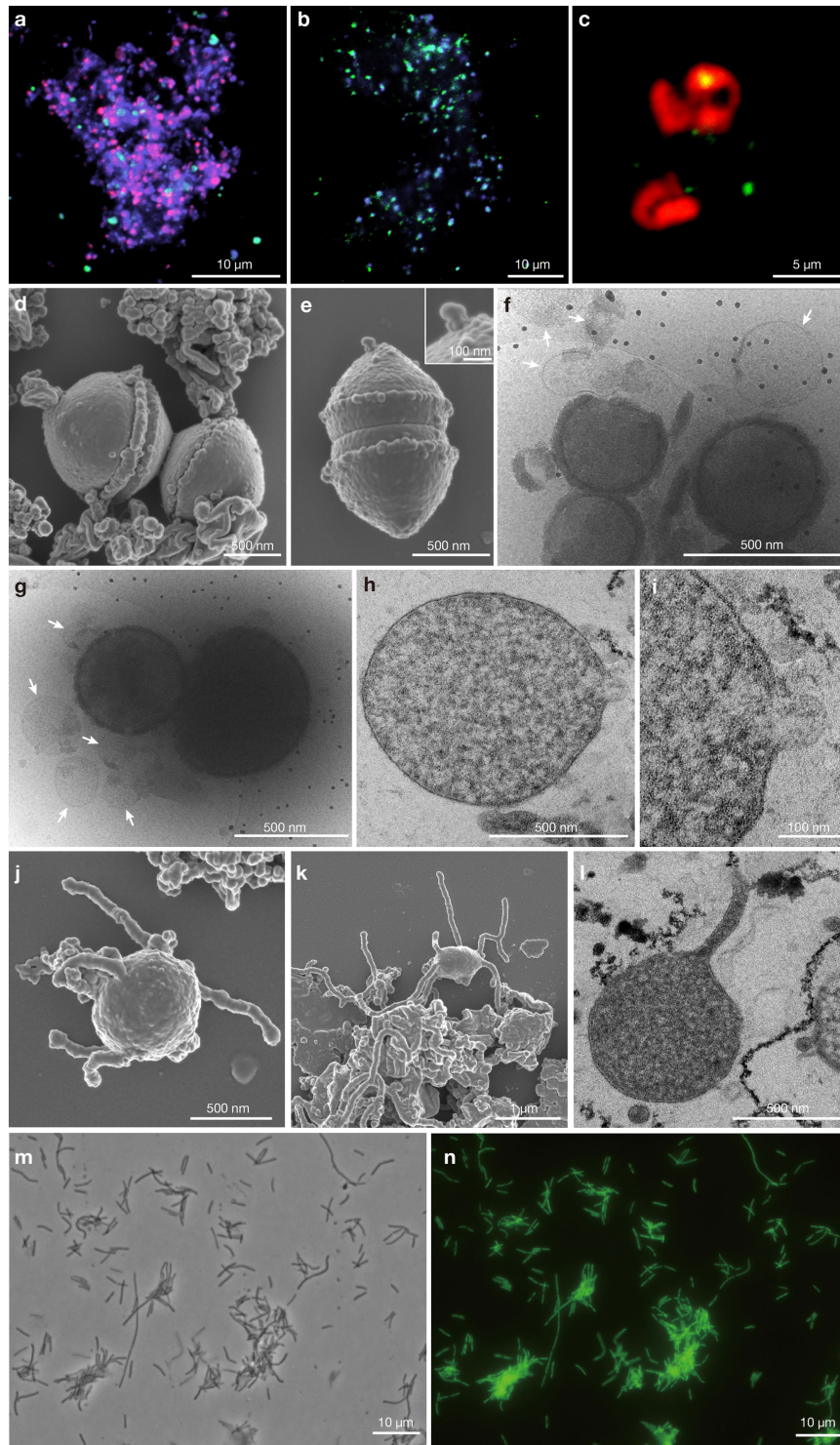
**Extended Data Fig. 1 | Growth of MK-D1. a**, Effect of temperature on growth of MK-D1. Data are mean  $\pm$  s.d. of triplicate determinations. Each data point is shown as a dot. The temperature range test was performed twice with similar results. **b, c**, The amino acid concentrations and growth curves of MK-D1 in pure cocultures at 20 °C. Results from cultures 1 (**b**) and 2 (**c**) are shown. Please note

that the initial concentrations of amino acids were normalized to 100%. Total amino acids and several representative amino acids (Val, valine; Leu, leucine; Ile, isoleucine) are independently shown for the duplicate culture samples. Detailed iTAG-based community compositions of the cultures are shown in Supplementary Table 1.



**Extended Data Fig. 2 | Circular representation of MK-D1 genome.** From the outside to the centre: the distribution of the coding sequences based on the conserved (orange) or non-conserved (grey) genes in the first circle, non-coding RNAs in the second circle, GC content showing deviation from average

(40.7%) in the third circle, and GC skew in the fourth circle. The GC content and GC skew were calculated using a sliding window of 2 kb in step of 10 kb. The coding sequences and RNA genes illustrate the findings for plus and minus strands.

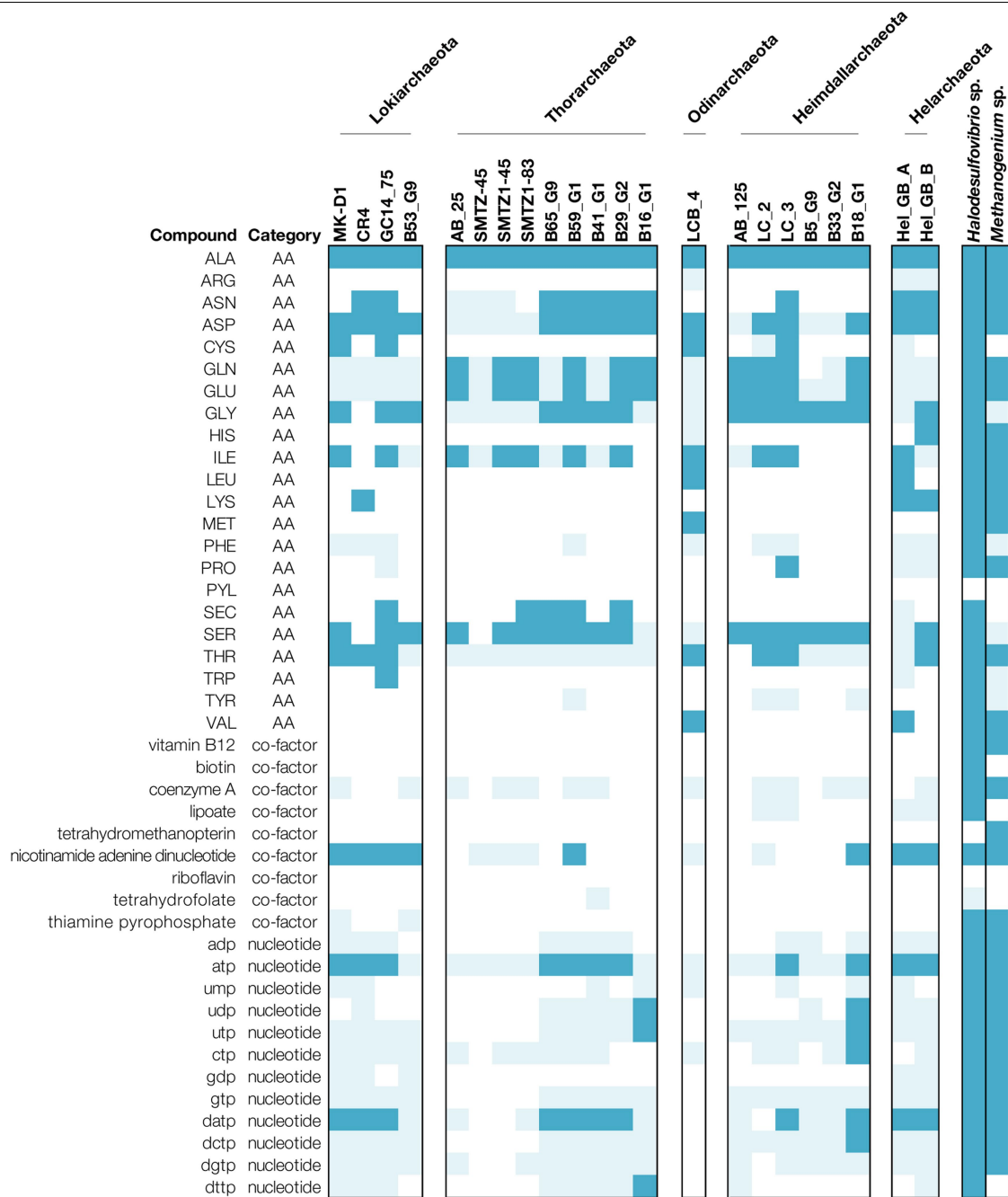


**Extended Data Fig. 3 | Other representative photomicrographs of MK-D1 cultures and *Methanobacterium* sp. strain MO-MB1.** **a, b,** Fluorescence images of cells from enrichment cultures after 8 (**a**) and 11 (**b**) transfers stained with DAPI (violet) and hybridized with nucleotide probes that target MK-D1 (green) and Bacteria (red). The images are different fields of view to those shown in Fig. 1b, c, which were taken at the same time. **c,** A fluorescence image of cells in the enrichments after 11 transfers hybridized with nucleotide probes that target MK-D1 (green) and Archaea (but with one mismatch against MK-D1; red). Large and irregular coccoid-shaped cells stained by only ARC915 are probably *Methanogenium*. **d, e,** Dividing cells of MK-D1 with a bleb. The top-right inset image in **e** shows a magnification of the bleb. **f, g,** Cryo-EM images of MK-D1 cells and large membrane vesicles (white arrows). **h, i, l,** Ultrathin sections of MK-D1 cells with a membrane vesicle. The image **i** shows a magnified image of **h, j, k,** SEM images of MK-D1 cells with protrusions. **l,** Ultrathin section of a

MK-D1 cell with a protrusion. **m, n,** Photomicrographs of pure culture of *Methanobacterium* sp. strain MO-MB1 cells stained with SYBR Green I. Phase-contrast (**m**) and fluorescence (**n**) images of the same field are shown. **a, b,** The FISH experiments were performed three times with similar results. **d, e, j, k,** The SEM images are representative of  $n = 122$  recorded images that were obtained from four independent observations from four culture samples. The lipid composition experiments were repeated twice and gave similar results. **f, g,** The cryo-EM images are representative of  $n = 14$  recorded images that were taken from two independent observations from two culture samples. **h, i, l,** The ultrathin-section images are representative of  $n = 131$  recorded images that were obtained from six independent observations from six culture samples. **m, n,** The SYBR Green I staining experiment was performed once, but all 10 recorded images showed similar results. Detailed iTAG analyses of cultures are shown in Supplementary Table 1.







**Extended Data Fig. 5 | Amino acid, cofactor and nucleotide biosynthesis capacities of MK-D1 and other Asgard archaea.** Genomes that encode proteins for the synthesis of amino acids, cofactors and nucleotides from pyruvate or acetyl-CoA (dark blue) and synthesis from other intermediates

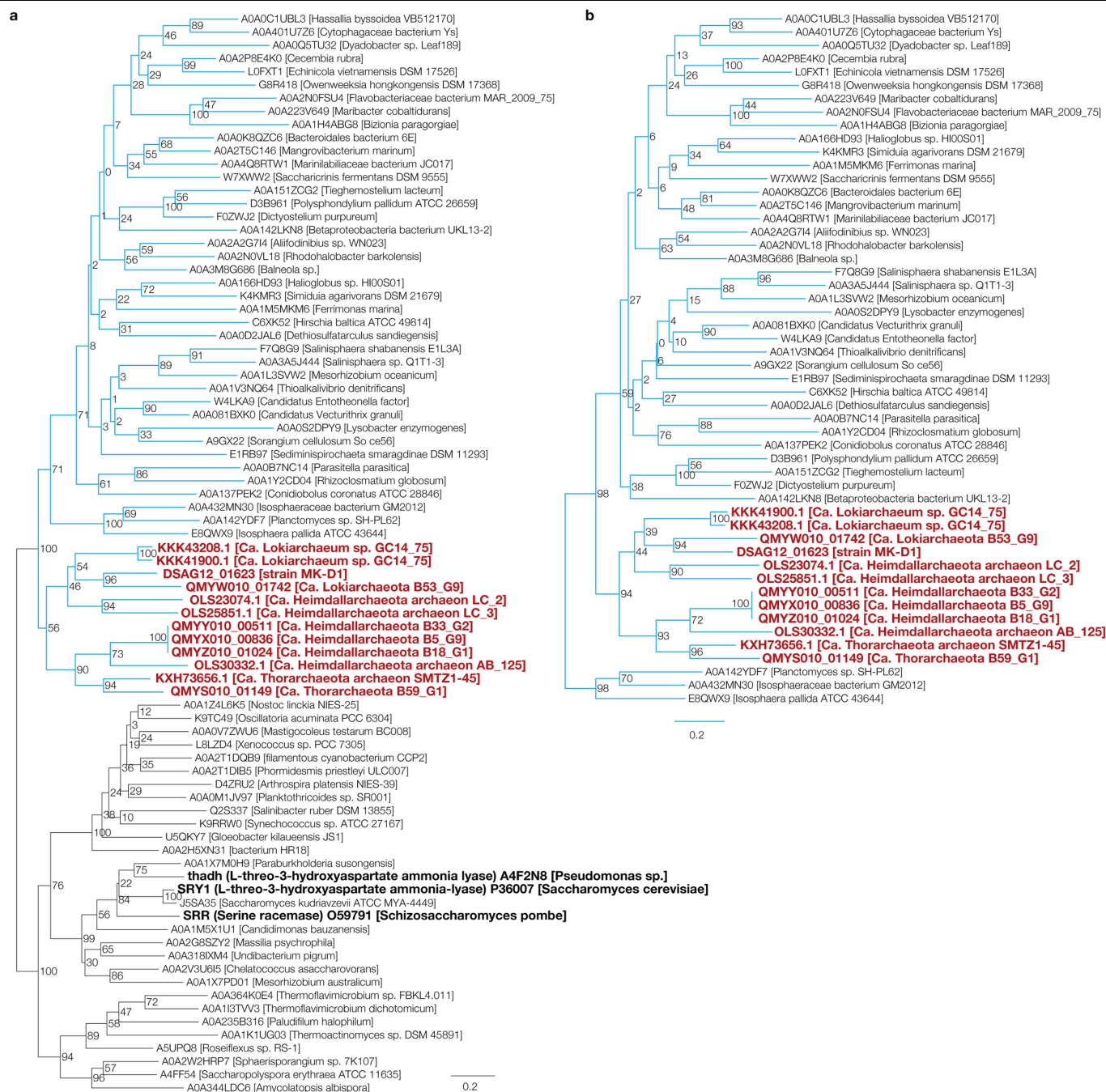
(light blue) are indicated. Those without complete pathways from pyruvate and/or acetyl-CoA are indicated in white. *Halodesulfobrio* sp. strain MK-HDV and *Methanogenium* sp. strain MK-MG isolated in this study are also shown.

## Urocanate hydratase



**Extended Data Fig. 6 | Maximum-likelihood tree of Asgard archaea urocanate hydratase.** Urocanate hydratase (HutU) homologues were obtained by BLASTp analysis of the Asgard archaea sequences against the UniProt database (release 2019\_06). Of homologues with sequence similarity  $\geq 40\%$  and overlap  $\geq 70\%$ , representative sequences were selected using CD-HIT with a clustering cut-off of 70% similarity (otherwise default settings were used). Additional homologues with verified biochemical activity, sequence similarity  $\geq 30\%$  and overlap  $\geq 70\%$  were obtained by BLASTp analysis of the Asgard

archaea sequences against the UniProt/SwissProt database (2019\_05). Sequences were aligned using MAFFT v.7 with default settings and trimmed using trimAl v.1.2 with default settings. The maximum-likelihood tree was constructed using RAxML-NG using fixed empirical substitution matrix (LG), 4 discrete GAMMA categories, empirical amino acid frequencies from the alignment and 100 bootstrap replicates. In total, 876 sites of the alignment were used for tree construction.



**Extended Data Fig. 7 | Maximum-likelihood tree of Asgard archaeal L-threonine/L-serine dehydratase. a**, Tree calculated for target Asgard archaeal L-threonine/L-serine dehydratase (TdcB) and homologues. TdcB homologues were obtained by BLASTp analysis of the Asgard archaea sequences against the UniProt reference proteome and SwissProt database (release 2019\_06). Of homologues with sequence similarity  $\geq 40\%$ , overlap  $\geq 70\%$  and predicted prosite domain PSO0165 (serine/threonine dehydratases pyridoxal-phosphate attachment site), representative sequences were selected using CD-HIT with a clustering cut-off of 70% similarity (otherwise default settings were used). Additional homologues with verified biochemical activity, sequence similarity  $\geq 30\%$  and overlap  $\geq 70\%$  were obtained by BLASTp analysis of the Asgard archaea sequences against the UniProt/SwissProt database (2019\_05).

Sequences were aligned using MAFFT v.7 with default settings. Positions with gaps in more than 10% of the sequences were excluded from the alignment using trimAl v.1.2 (-gt 0.9; and otherwise default settings were used). The maximum-likelihood tree was constructed using PhyML using a fixed empirical substitution matrix (LG), 4 discrete GAMMA categories, empirical amino acid frequencies from the alignment and 100 bootstrap replicates (-b 100 -d aa -m LG -v e). In total, 308 sites of the alignment were used for tree construction.

**b**, Tree calculated for a subset of sequences contained in a section of the original tree (branches that are coloured blue). Sequences were realigned and trimmed as described for **a**. In total, 308 sites of the alignment were used for tree construction.

## Extended Data Table 1 | SSU rRNA gene clones obtained from the primary and six successive transferred enrichment cultures

### <Primary enrichment culture>

#### Clone library using universal primers (530F/907R)

Phylotype name	No. of clones	Accession no.	Sequence length (bp)	Closest cultured species or clone (accession number)	Sequence identity (%)	Phylogenetic affiliation	Identical or almost identical clones detected from the AOM bioreactor enrichment (accession number, sequence identity %) <sup>a</sup>
111_U1	40	LC490621	374	<i>Halodesulfobivrio aestuarii</i> strain Sylt 3 (NR_116770)	99	genus <i>Halodesulfobivrio</i>	—
111_U2	3	—	377	<i>Methylobacter marinus</i> strain A45 (NR_025132)	100	genus <i>Methylobacter</i>	MK903D_B19 (AB831411, 100%)
111_U3	2	—	374	<i>Photobacterium indicum</i> strain NBRC 14233 (NR_113657)	100	genus <i>Photobacterium</i>	MK903D_B9 (AB831402, 100%)
111_U4	1	LC490622	377	subseafloor sediment clone ODP1251B13.14 (AB177314)	99	subgroup 21 within the phylum <i>Acidobacteria</i>	—
111_U5	1	LC490623	377	hydrothermal seep sediment BAC_OTU_13 (KP091106)	100	GIF9 group within the class <i>Dehalococcoidia</i>	MK0D_B60 (AB831337, 99.5%)
111_U6	1	LC490624	374	<i>Roseovarius gaetbuli</i> strain YM-20 (NR_134163)	99	genus <i>Roseovarius</i>	—

#### Clone library using archaeal primers (340F/932R)

Phylotype name	No. of clones	Accession no.	Sequence length (bp)	Closest cultured species or clone (accession number)	Sequence identity (%)	Phylogenetic affiliation	Identical or almost identical clones detected from the AOM bioreactor enrichment (accession number, sequence identity %) <sup>a</sup>
111_A1	6	—	535	<i>Methanococcoides burtonii</i> strain DSM 6242 (NR_074242)	99	genus <i>Methanococcoides</i>	MK903D_A2 (AB831282, 100%)
111_A2	5	LC490620	513	<i>Methanogenium cariaci</i> strain JR1 (NR_104730)	99	genus <i>Methanogenium</i>	—
111_A3	2	—	534	methane seep clone AN_5119N_arc_E4_T3 (KM356859)	99	ANME-2a	MK0D_A9 (AB831268, 100%)
111_A4	2	LC490619	516	methane seep clone AC_5120_arc_D2_T3 (KM356804)	99	Lokiarchaeota ( <i>Ca. P. syntrophicum</i> strain MK-D1)	MK903R_A35 (AB831305, 99.0%)

### <Six successive transferred enrichment culture>

#### Clone library using universal primers (530F/907R)

Phylotype name	No. of clones	Accession no.	Sequence length (bp)	Closest cultured species or clone (accession number)	Sequence identity (%)	Phylogenetic affiliation	Identical or almost identical clones detected from the AOM bioreactor enrichment (accession number, sequence identity %) <sup>a</sup>
111-5_U1	40	—	374	<i>Halodesulfobivrio oceani</i> strain I.8.1 (NR_116768)	100	genus <i>Halodesulfobivrio</i>	—
111-5_U2	6	—	380	methane seep clone AC_5120_arc_D2_T3 (KM356804)	100	Lokiarchaeota ( <i>Ca. P. syntrophicum</i> strain MK-D1)	—
111-5_U3	1	—	380	<i>Methanogenium boonei</i> strain AK-7 (NR_115706)	99	genus <i>Methanogenium</i>	—

<sup>a</sup>The clone sequences have been reported in our previous study<sup>15</sup>.



Extended Data Table 2 | Carbon isotope fractionation values in MK-D1 cultures after 120 days incubation with and without stable isotope labelled amino acids

Culture ID	$\delta^{13}\text{C}\text{-CO}_2$ (‰ VPDB) <sup>a</sup>	$\delta^{13}\text{C}\text{-CH}_4$ (‰ VPDB) <sup>a</sup>
<b><i>Co-cultures with Methanobacterium</i></b>		
No.1 with stable isotope labeled AAs	-12.3	4094.8
No.2 with stable isotope labeled AAs	-9.3	6990.7
No.3 w/o stable isotope labeled AAs	-23.1	-36.7
No.4 w/o stable isotope labeled AAs	-23.1	-33.1
<b><i>Tri-cultures with Halodesulfovibrio and Methanogenium</i></b>		
No.5 with stable isotope labeled AAs	318.5	86.0
No.6 with stable isotope labeled AAs	309.3	87.8
No.7 w/o stable isotope labeled AAs	-22.6	-95.5
No.8 w/o stable isotope labeled AAs	-22.8	-97.8

<sup>a</sup>Parts per thousand (‰) compared with the Vienna Pee Dee Belemnite (VPDB) standard.

**Extended Data Table 3 | Growth of MK-D1 after incubation of 120 days with a range of substrates**

Culture name	Substrate	Initial MK-D1 16S rRNA gene copies per ml of culture	Final MK-D1 16S rRNA gene copies per ml of culture	No. of MK-D1 16S rRNA gene copies relative to initial culture	Community compositions evaluated by iTAG analysis (%) <sup>a</sup>			
					MK-D1	<i>Methanogenium</i>	<i>Methenobacterium</i> sp. strain MO-MB1	Others
Inoculum	Casamino acids (CA) <sup>b</sup> + 20 amino acids mixture (AAs) <sup>c</sup> + powdered milk (PM) <sup>d</sup>	—	5.91E+05	—	39.8	36.8	23.3	0.01
Control-1	CA + 20 AAs + PM	1.42E+04	1.62E+05	11.36	76.7	21.8	1.4	0.03
Control-2	CA + 20 AAs + PM	4.67E+03	6.55E+04	14.03	60.3	38.0	1.6	0.04
H2-1	CA + 20 AAs + PM + 1.5 kPa H <sub>2</sub> <sup>e</sup> + 10 mM 2-bromoethane sulfonate (2-BES) <sup>f</sup>	9.46E+03	4.35E+03	0.46	—	—	—	—
H2-2	CA + 20 AAs + PM + 1.5 kPa H <sub>2</sub> + 10 mM 2-BES	1.37E+04	3.28E+03	0.24	—	—	—	—
H2-3	CA + 20 AAs + PM + 1.5 kPa H <sub>2</sub> + 10 mM 2-BES	3.10E+04	8.27E+03	0.27	—	—	—	—
Formate-1	CA + 20 AAs + PM + 1 mM Formate + 10 mM 2-BES	2.76E+04	2.00E+03	0.07	—	—	—	—
Formate-2	CA + 20 AAs + PM + 1 mM Formate + 10 mM 2-BES	1.46E+04	9.49E+03	0.65	—	—	—	—
Nitrate-1	CA + 20 AAs + PM + 500 $\mu$ M Nitrate <sup>g</sup>	2.13E+04	8.43E+03	0.40	—	—	—	—
Nitrate-2	CA + 20 AAs + PM + 500 $\mu$ M Nitrate	1.47E+04	5.19E+03	0.35	—	—	—	—
Sulfate-1	CA + 20 AAs + PM + 500 $\mu$ M Sulfate	5.28E+03	9.21E+04	17.42	79.5	19.5	1.0	0.03
Sulfate-2	CA + 20 AAs + PM + 500 $\mu$ M Sulfate	3.39E+04	5.28E+04	1.56	—	—	—	—
Thiosulfate-1	CA + 20 AAs + PM + 500 $\mu$ M Thiosulfate	1.23E+04	5.00E+04	4.05	—	—	—	—
Thiosulfate-2	CA + 20 AAs + PM + 500 $\mu$ M Thiosulfate	2.29E+04	6.09E+04	2.66	—	—	—	—
Lactate-1	CA + 20 AAs + PM + 1 mM Lactate	5.31E+03	1.31E+04	2.46	—	—	—	—
Lactate-2	CA + 20 AAs + PM + 1 mM Lactate	1.53E+04	1.91E+04	1.25	—	—	—	—
Acetate-1	CA + 20 AAs + PM + 1 mM Acetate	2.63E+04	9.17E+04	3.48	—	—	—	—
Acetate-2	CA + 20 AAs + PM + 1 mM Acetate	1.56E+04	2.13E+04	1.36	—	—	—	—
Glucose-1	CA + 20 AAs + PM + 1 mM Glucose	1.12E+04	1.16E+05	10.33	73.8	24.3	1.9	0.03
Glucose-2	CA + 20 AAs + PM + 1 mM Glucose	1.06E+04	1.06E+05	10.01	70.3	28.0	1.7	Not detected
Fructose-1	CA + 20 AAs + PM + 1 mM Fructose	3.18E+04	3.31E+04	1.04	—	—	—	—
Fructose-2	CA + 20 AAs + PM + 1 mM Fructose	1.79E+04	1.44E+05	8.08	—	—	—	—
Xylose-1	CA + 20 AAs + PM + 1 mM Xylose	2.82E+04	6.79E+03	0.24	—	—	—	—
Xylose-2	CA + 20 AAs + PM + 1 mM Xylose	9.25E+03	1.18E+05	12.73	61.4	36.5	2.1	0.01
Ribose-1	CA + 20 AAs + PM + 1 mM Ribose	1.42E+04	2.88E+04	2.02	—	—	—	—
Ribose-2	CA + 20 AAs + PM + 1 mM Ribose	7.34E+03	2.29E+04	3.13	—	—	—	—
Maltose-1	CA + 20 AAs + PM + 1 mM Maltose	2.84E+04	1.21E+05	4.25	—	—	—	—
Maltose-2	CA + 20 AAs + PM + 1 mM Maltose	2.17E+04	4.55E+04	2.09	—	—	—	—
Citrate-1	CA + 20 AAs + PM + 1 mM Citrate	3.36E+04	1.20E+05	3.56	—	—	—	—
Citrate-2	CA + 20 AAs + PM + 1 mM Citrate	1.82E+04	5.73E+04	3.15	—	—	—	—
Pyruvate-1	CA + 20 AAs + PM + 1 mM Pyruvate	1.73E+04	9.37E+04	5.42	—	—	—	—
Pyruvate-2	CA + 20 AAs + PM + 1 mM Pyruvate	2.22E+04	4.86E+03	0.22	—	—	—	—
Fumarate-1	CA + 20 AAs + PM + 1 mM Fumarate	3.16E+04	7.20E+04	2.28	—	—	—	—
Fumarate-2	CA + 20 AAs + PM + 1 mM Fumarate	1.94E+04	2.35E+04	1.21	—	—	—	—
Archaeal cell-1	CA + 20 AAs + PM + archaeal cell membrane components <sup>h</sup>	1.53E+04	1.42E+05	9.27	81.5	17.5	0.8	0.3
Archaeal cell-2	CA + 20 AAs + PM + archaeal cell membrane components	4.17E+04	1.05E+05	2.52	—	—	—	—

A dash indicates that data were not taken for that sample. <sup>a</sup>The iTAG analysis was performed for samples in which an increase of about 10 times or more in 16S rRNA gene copy numbers of MK-D1 was observed after incubation; data were analysed by qPCR assay. Detailed results are shown in Supplementary Table 1. <sup>b</sup>Final concentration of casamino acids was 0.05% (w/v). <sup>c</sup>Final concentration of each amino acid was 0.1 mM. <sup>d</sup>Powdered milk for baby (Hohoemi, Meiji) was used at a final concentration of 0.1% (w/v). <sup>e</sup>The concentration of hydrogen gas was in the head space of the culture bottle. <sup>f</sup>2-BES was added to inhibit methanogens. <sup>g</sup>Addition of nitrate completely suppressed the growth of MK-D1. This is probably because nitrate inhibits formate dehydrogenase activity of MK-D1<sup>95</sup>. <sup>h</sup>Archaeal cell membrane components were a mixture of phytol, intact polar lipid–glycerol–dialkyl–glycerol tetraethers and core lipid– glycerol–dialkyl–glycerol tetraethers (each at a final concentration 50 ng ml<sup>-1</sup>). We used the archaeal membrane components as these have a positive effect on the growth of some archaeal species: (i) archaeal cell extract including membrane lipids stimulates the growth of the extremely thermophilic archaeon *Thermocodium modestius*<sup>96</sup>, and (ii) the hyperthermophilic archaeon *Thermofilum pendes* requires the polar lipids for growth, which was obtained from the archaeal species *Thermoproteus tenax*<sup>97</sup>.

## Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- |                                     |   |
|-------------------------------------|---|
| n/a                                 | Confirmed   |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement   |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided<br><i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A description of all covariates tested   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons  |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For null hypothesis testing, the test statistic (e.g. $F$ , $t$ , $r$ ) with confidence intervals, effect sizes, degrees of freedom and $P$ value noted<br><i>Give <math>P</math> values as exact values whenever suitable.</i>                            |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Estimates of effect sizes (e.g. Cohen's $d$ , Pearson's $r$ ), indicating how they were calculated   |

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

### Software and code

Policy information about [availability of computer code](#)

Data collection	MiSeq control software v3.1 for sequencing; StepOne Software v2.3 for qPCR; Nikon NIS-Elements software for FISH imaging
Data analysis	IDL based NASA JSC imaging software, OpenMIMS ( <a href="https://github.com/BWHCNI/OpenMIMS">https://github.com/BWHCNI/OpenMIMS</a> ), and Look@NanoSIMS for NanoSIMS; IMOD for tomogram reconstruction; Trimmomatic v0.33 and NextClip v1.3.1 for trimming; SPAdes v3.1.1 for genome assembly; MyCC (2015/07/10) for genome binning; SSPACE v3.0 for scaffolding; Prokka v1.12 for genome annotation; SignalP v4.1 for signal peptide prediction; MAFFT v7 for gene sequence alignment; RAXML-NG v0.8.0 for maximum likelihood tree construction; trimAl v1.2 for alignment trimming; CD-HIT v4.8.1 for gene clustering; MrBayes 3.2.7a for Bayesian phylogenetic tree calculation; PhyML 3.3 for maximum likelihood tree construction; MUSCLE v3.8.31 for sequence alignment.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

### Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Genomes for *Ca. Prometheoarchaeum syntrophicum* MK-D1, *Halodesulfovibrio* sp. MK-HDV, and *Methanogenium* sp. MK-MG are available under Genbank BioProjects PRJNA557562, PRJNA557563, and PRJNA557565 respectively. The iTAG sequence data was deposited in Bioproject PRJDB8518 with the accession numbers DRR184081–DRR184101. The 16S rRNA gene sequences of MK-D1, *Halodesulfovibrio*, *Methanogenium* and clones obtained from primary enrichment culture were deposited in the DDBJ/EMBL/GenBank database under accession numbers LC490619–LC490624.

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences    ☐ Behavioural & social sciences    ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample size-based calculations not relevant to analyses in this study
Data exclusions	No data were excluded from the analysis
Replication	Culture experiments were performed in duplicate or triplicate. RNA-based experiments were performed without replicates due to challenges in cultivation (i.e., extremely low growth rates and culture densities)
Randomization	Randomization not relevant to data collection/analyses in this study as the study does not involve participant groups. Each experiment included controls.
Blinding	Blinding not relevant to data collection performed in this study as blinding is not required and was not possible for cultivation-based experiments as the investigators must verify the control and non-control groups for each experiment.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data

### Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

# Nearest neighbours reveal fast and slow components of motor learning

<https://doi.org/10.1038/s41586-019-1892-x>

Sepp Kollmorgen<sup>1\*</sup>, Richard H. R. Hahnloser<sup>1</sup> & Valerio Mante<sup>1\*</sup>

Received: 25 January 2019

Accepted: 4 November 2019

Published online: 8 January 2020

Changes in behaviour resulting from environmental influences, development and learning<sup>1–5</sup> are commonly quantified on the basis of a few hand-picked features<sup>2–4,6,7</sup> (for example, the average pitch of acoustic vocalizations<sup>3</sup>), assuming discrete classes of behaviours (such as distinct vocal syllables)<sup>2,3,8–10</sup>. However, such methods generalize poorly across different behaviours and model systems and may miss important components of change. Here we present a more-general account of behavioural change that is based on nearest-neighbour statistics<sup>11–13</sup>, and apply it to song development in a songbird, the zebra finch<sup>3</sup>. First, we introduce the concept of ‘repertoire dating’, whereby each rendition of a behaviour (for example, each vocalization) is assigned a repertoire time, reflecting when similar renditions were typical in the behavioural repertoire. Repertoire time isolates the components of vocal variability that are congruent with long-term changes due to vocal learning and development, and stratifies the behavioural repertoire into ‘regressions’, ‘anticipations’ and ‘typical renditions’. Second, we obtain a holistic, yet low-dimensional, description of vocal change in terms of a stratified ‘behavioural trajectory’, revealing numerous previously unrecognized components of behavioural change on fast and slow timescales, as well as distinct patterns of overnight consolidation<sup>1,2,4,14,15</sup> across the behavioral repertoire. We find that diurnal changes in regressions undergo only weak consolidation, whereas anticipations and typical renditions consolidate fully. Because of its generality, our nonparametric description of how behaviour evolves relative to itself—rather than to a potentially arbitrary, experimenter-defined goal<sup>2,3,14,16</sup>—appears well suited for comparing learning and change across behaviours and species<sup>17,18</sup>, as well as biological and artificial systems<sup>5</sup>.

Zebra finches acquire complex, stereotyped vocalizations through a months-long process of sensory–motor learning<sup>3,19–21</sup>. During development, syllable order—that is, syntax—and the spectral structure of syllables evolve<sup>3</sup>. These two aspects of vocal learning may be mediated by largely independent mechanisms with distinct anatomical substrates<sup>21,22</sup>. Here we focus on characterizing the development of spectral structure. We began our studies by obtaining dense audio recordings of five male zebra finches between 35 and 123 days post-hatch (dph; mean  $\pm$  standard deviation  $73.4 \pm 18.6$  consecutive days of recording). Birds were isolated from other males after birth and, on average, live-tutored from around 46 to 63 dph (Extended Data Fig. 1a). Band-passed (0.35–8 kHz) audio recordings were segmented into individual vocal renditions, and represented as song spectrogram segments (Fig. 1a; 563,124–1,203,647 renditions per bird). We excluded noise and isolated calls from the analyses.

## Behavioural change in single features

Vocal development is often characterized by considering changes in acoustic features such as pitch, frequency modulation<sup>3</sup> or entropy variance<sup>2,14</sup> (Fig. 1b). Such characterizations readily reveal multiple

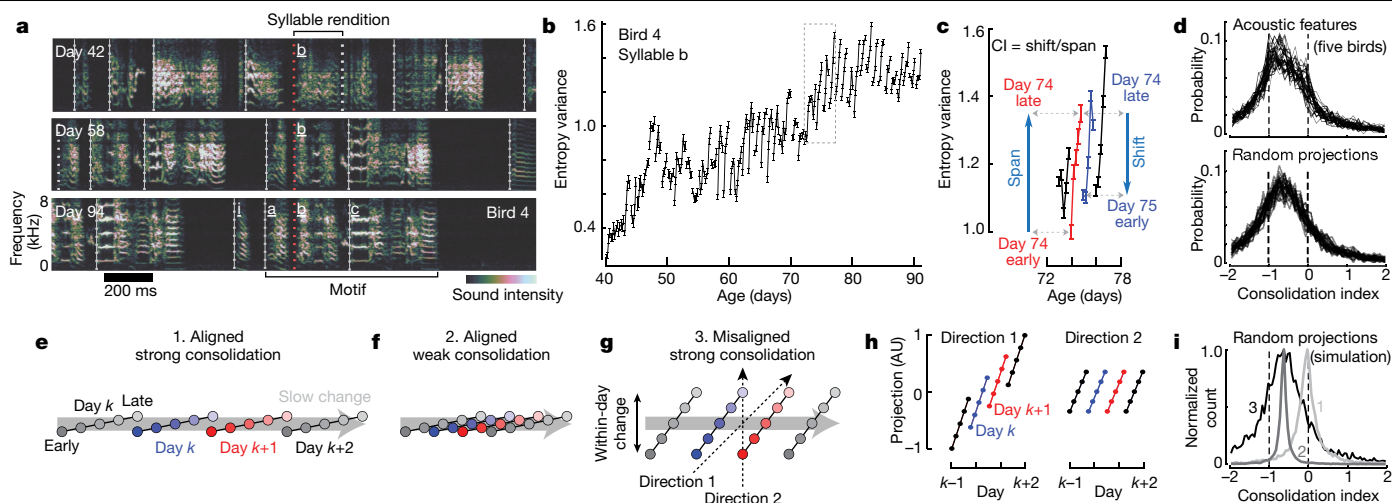
timescales of behavioural change: individual features can vary consistently within a day, display overnight discontinuities, and show drift over the duration of weeks or months (Fig. 1b, c).

We summarize the relation between change at these different timescales through a consolidation index (Fig. 1c), which measures whether within-day change in a feature (‘span’, Fig. 1c) is maintained or lost overnight (‘shift’, Fig. 1c). Weak consolidation<sup>2,14</sup> corresponds to a consolidation index of close to  $-1$  (no consolidation: the shift is equal but opposite to span); strong consolidation<sup>4,15</sup> corresponds to an index of close to  $0$  (perfect consolidation: the shift is  $0$  days); and offline learning<sup>4,23,24</sup> to an index of larger than  $0$ . Across 32 commonly used acoustic features, the consolidation indices in our data are mostly negative, indicating weak consolidation (Fig. 1d, top; median  $-0.67$ ). This finding holds even for random spectral features (Fig. 1d, bottom; median  $-0.64$ ) and is consistent with past accounts of song development in zebra finches<sup>2,14</sup>.

Individual features, however, may provide an incomplete account of change in a complex behaviour such as song vocalizations. To illustrate this point, we consider three simple scenarios. In the first two (Fig. 1e, f), the change in behaviour that occurs within any given day largely mirrors, on a faster timescale, the slow change that occurs over the course

<sup>1</sup>Institute of Neuroinformatics and Neuroscience Center Zurich, University of Zurich and ETH Zurich, Zurich, Switzerland. \*e-mail: skollmor@ini.ethz.ch; valerio@ini.uzh.ch





**Fig. 1 | Fast and slow change in developing zebra finch vocalizations.** **a**, Vocalizations at three developmental stages. Dotted lines indicate syllable onsets. Crystallized song syllables (middle and bottom) fall into discrete categories (syllables i, a, b, c) and form a stereotyped ‘motif’, typically resembling the tutor song. **b**, Time course of one acoustic feature, entropy variance, for syllable b. **c**, Magnification of the region outlined in **b**, showing a period of within-day span (early to late, day  $k$ ) and overnight shift (late day  $k$  to early day  $k+1$ ). The consolidation index (CI) is approximately  $-0.75$ . **d**, Histograms of consolidation indices over pairs of consecutive days, syllables

and birds, for 32 acoustic features (top) and 32 random spectral projections (bottom). **e–g**, Three scenarios of slow developmental change (grey arrows) and fast within-day change in vocalizations. Each point represents the distribution of vocalizations from a given time and day. A larger distance between points indicates more dissimilar distributions. **h**, Linear projections of the points in **g** onto two example song features (dotted lines in **g**) for the misaligned, strong-consolidation scenario. Consolidation strength varies across directions. **i**, Consolidation indices over 10,000 random projections simulated from the three scenarios (1, 2 and 3 in **e–g**).

of many days or weeks. In the third scenario (Fig. 1g), within-day change is partly ‘misaligned’ with slow change: that is, it involves behavioural features that do not consistently change on slower timescales. Within-day change could reflect metabolic, neural or other changes that are not necessarily congruent with longer-term learning or development; the slow change reflects long-term modifications in behaviour that are typically equated with learning and development. We abstractly refer to these slow components as the direction of slow change (DiSC).

Notably, simulations of these scenarios show that negative consolidation indices for single features can result from very different time courses of development (Fig. 1h, i). Negative indices occur both when within-day and slow changes are closely aligned but daily gains along the DiSC are mostly lost overnight (weak consolidation, Fig. 1f), and when diurnal gains along the DiSC are perfectly consolidated but within-day change is substantially misaligned with slow change (Fig. 1g). The broad distributions of indices observed during song development (Fig. 1d, top), which also include strongly positive indices, seem more consistent with the misaligned scenario (Fig. 1i, histogram 3).

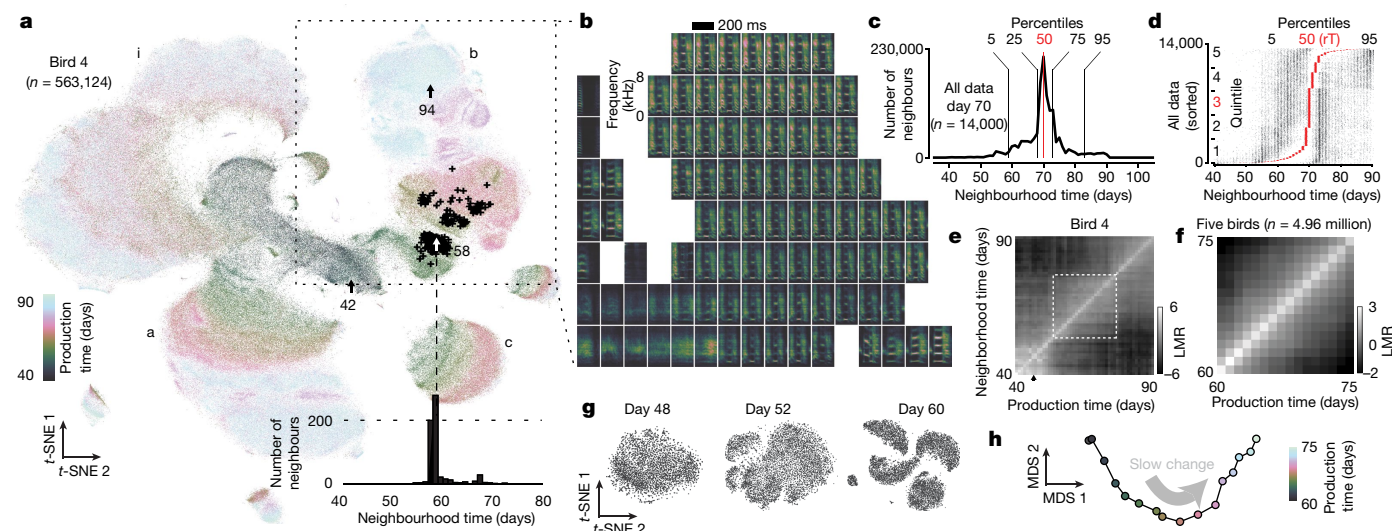
## Nearest-neighbour measures of change

We developed a general characterization of change in high-dimensional behavioural data, based on nearest-neighbour statistics<sup>12,13</sup>, that can distinguish between the scenarios in Fig. 1e–g. We initially analyse song-spectrogram segments of fixed duration aligned to syllable onset (Fig. 1a), but later extend our analysis to alternative parameterizations of the vocalization behaviour. Vocal renditions are represented as real-valued vectors  $x_i \in \mathbb{R}^d$  (where  $i$  indexes renditions, and  $d$  denotes dimension), each associated with a production time,  $t_i \in \mathbb{R}$  (for example, the bird’s age when singing  $x_i$ ). The  $K$ -neighbourhood of rendition  $x_i$  is given by those  $K$  renditions (among the set of all renditions) that are closest to  $x_i$  on the basis of some metric (for example, Euclidean distance). For small-enough values of  $K$ , different syllable types do not mix within a neighbourhood (Extended Data Fig. 1e) and neighbourhood statistics are largely independent of cluster boundaries, obviating the need for clustering renditions into syllables.

We visualize all vocalizations produced by a bird throughout development with Barnes–Hut  $t$ -distributed stochastic neighbour embedding ( $t$ -SNE)<sup>11</sup> (which predominantly preserves local neighbourhoods<sup>11</sup>). Each point in the embedding corresponds to a spectrogram segment,  $x_i$  (Fig. 1a). Different locations correspond to different vocalization types (Fig. 2b and Extended Data Fig. 2a). The embedding suggests that vocalizations change from undifferentiated subsong<sup>3,20</sup> (Fig. 2a, middle) to clearly differentiated syllables that fall into at least four categories (Fig. 2a, syllables a, b, c and introductory note i, as in Fig. 1a). The emergence of clustered syllables from unclustered subsong can be confirmed by standard clustering approaches (Fig. 2g and Extended Data Fig. 1c, d). Notably, the embedding does not preserve all local structure in the data, as nearest neighbours in the embedding space are not necessarily nearest neighbours in the high-dimensional data space (Fig. 2a; black crosses represent high-dimensional neighbours). We therefore quantify behavioural change directly in the high-dimensional data by analysing the composition of high-dimensional neighbourhoods<sup>12,13</sup> (Extended Data Fig. 2e–g).

For each data point, we refer to the production times of all data points in its  $K$ -neighbourhood as ‘neighbourhood production times’ (or ‘neighbourhood times’; Fig. 2a, histogram). We summarize the neighbourhood times of many data points (Fig. 2d) through ‘pooled neighbourhood times’ (Fig. 2c) and the ‘neighbourhood mixing matrix’ (Fig. 2e and Extended Data Figs. 2g, 3d). Each value in the neighbourhood mixing matrix represents the similarity between behaviours from two production periods. Deviations from zero indicate that behaviours from the corresponding production periods are more similar (for values greater than 0), in terms of mixing at the level of  $K$ -neighbourhoods, or less similar (for values smaller than 0) than expected from a shuffling null hypothesis.

We use multidimensional scaling<sup>25</sup> on the mixing matrix to represent the similarity between behaviours from different production times as a ‘behavioural trajectory’ (Fig. 2h). Each point on the trajectory represents the distribution of all vocalizations produced on a given day. Pairwise distances between points represent the dissimilarity between distributions (Extended Data Fig. 2e–g). Here we focus on a 16-day



**Fig. 2 | Neighbourhood mixing and repertoire dating.** **a**, *t*-SNE of all vocalizations from the bird in Fig. 1a. Each point is a syllable rendition. Clusters (syllables i, a, b, c) emerge during development. Arrows indicate renditions of syllable b from Fig. 1a. Crosses show the 600 nearest neighbours of the rendition from day 58. Inset, histogram of production times (neighbourhood times) over the 600 nearest neighbours. **b**, Average spectrograms for different locations in the *t*-SNE visualization from **a**. **c**, Pooled neighbourhood times for day 70. Percentiles (vertical lines) quantify the extent of the behavioural repertoire on day 70. **d**, Percentiles (5th, 50th and 95th) of neighbourhood times for individual renditions from day 70 (each row represents a rendition). Rows are sorted by the 50th percentile—the repertoire time (rT, red dots). Left

and right black dots mark the 5th and 95th percentiles. A small random horizontal shift was added to each dot for visualization. **e**, Mixing matrix for all data points depicted in **a**. Each column of the matrix represents a histogram of production times, pooled over all neighbourhoods of points within a day (*x* axis), normalized by a shuffling null hypothesis (LMR, base-2 logarithm of the mixing ratio). The black arrow marks the first day of tutoring. **f**, Average mixing matrix for five birds (days 60–75). **g**, Single-day *t*-SNE, for three days (for the same bird as in Fig. 1a), illustrating the gradual emergence of clusters. **h**, Behavioural trajectory based on **f**, computed with ten-dimensional multidimensional scaling (MDS). Each point corresponds to a day. The two dimensions that capture the most variance in the trajectory are shown.

phase of gradual change midway through development (Fig. 2f). During this phase, the behavioural trajectory is structured differently on fast and slow timescales (Extended Data Fig. 3f–h). The two-dimensional projection of the trajectory that explains the maximal variance mainly reflects the direction of slow change (Fig. 1e–g, 2h).

The behavioural trajectory summarizes the progressive differentiation of vocalizations into distinct syllables, as well as simultaneous, continuous change in many spectral features of individual syllables. Notably, change is characterized through the behavioural trajectory by comparing the bird's song to itself across time, rather than to a tutor song. Thus the behavioural trajectory may also reflect innate song priors that can result in crystallized song deviating from the tutor song<sup>26</sup> and additional change due to other developmental processes<sup>27</sup>.

## Repertoire extent and consolidation

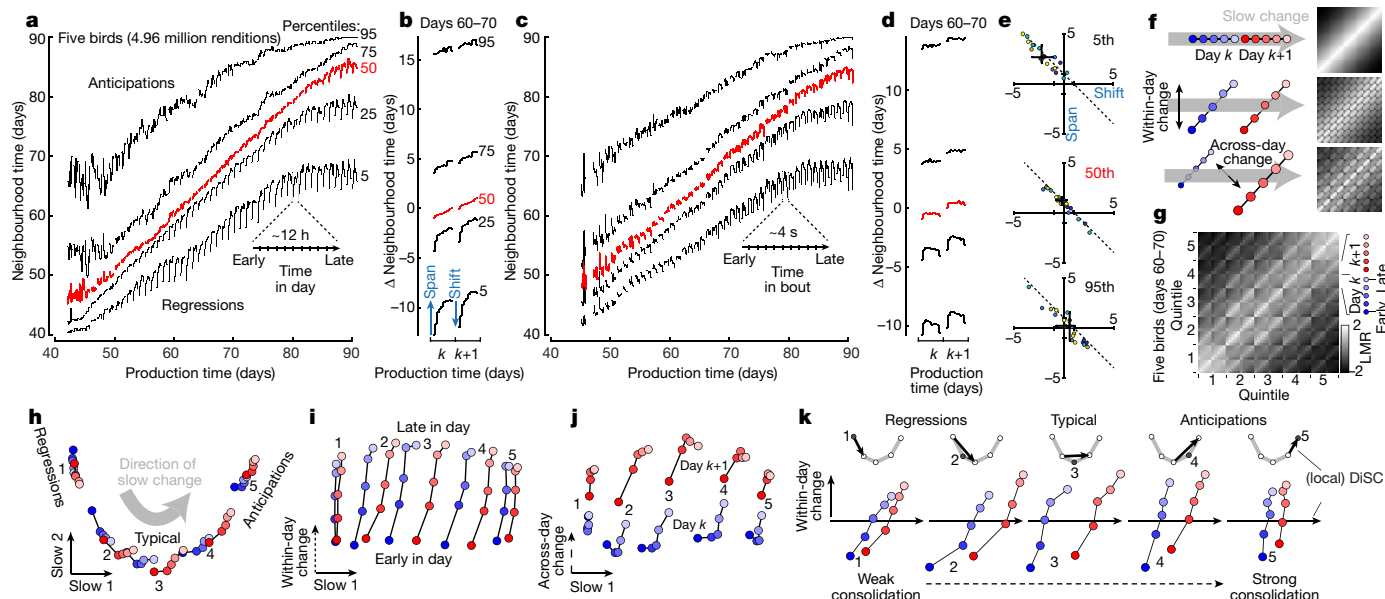
Additional *t*-SNE visualizations of the data suggest that renditions from nearby days overlap considerably, whereby changes occurring within a day partly mimic the slow change across days (Extended Data Fig. 2b, c). We quantify this apparent spread along the DiSC—reflecting different degrees of behavioural ‘maturity’—through neighbourhood times (Fig. 2d). We refer to behavioural renditions that predominantly have neighbours produced in the future as ‘anticipations’, and to renditions that predominantly have neighbours that were produced in the past as ‘regressions’ (Extended Data Fig. 3b). By contrast, renditions that are ‘typical’ for a given developmental stage mostly have neighbours produced on the same or nearby days. We denote the median neighbourhood time as the ‘repertoire time’ of a rendition. The repertoire time effectively places each rendition along the DiSC (Fig. 2d, *x* axis): that is, it dates it with respect to the progression of vocal development (‘repertoire dating’). A broad distribution of repertoire times across all renditions in a day (Fig. 2d) suggests considerable behavioural variability along the DiSC; the most extreme regressions are backdated

more than ten days into the past, and the most extreme anticipations are post-dated more than ten days into the future.

To quantify behavioural change on the timescale of hours, we subdivide each day into ten consecutive periods, and compute pooled neighbourhood times separately for each period. The percentiles of the pooled neighbourhood times chart the evolution of behaviour within and across days throughout development (Fig. 3a). Each repertoire-dating percentile is akin to a learning curve for a part of the behavioural repertoire (for example, typical renditions are described by the 50th percentile, and extreme anticipations by the 95th). The evolution of each percentile captures the progress along the DiSC (Fig. 3a, *y* axis) over time (Fig. 3a, *x* axis). We validated this characterization of behavioural change on simulated behaviour that mimicked vocal development (Extended Data Fig. 4a–d).

The repertoire-dating percentiles reveal that typical renditions move gradually along the DiSC throughout the day, and that changes along the DiSC acquired during the day are, on average, fully consolidated overnight (Fig. 3a, b, red). Anticipations undergo a similar or smaller degree of within-day change (Fig. 3a, b, 75th and 95th percentiles), whereas regressions move by a larger distance within each day, but this change is only weakly consolidated overnight (Fig. 3a, b, 5th and 25th percentiles; Fig. 3e). The most ‘immature’ renditions thus improve markedly throughout a day—more than typical renditions or anticipations—but these improvements are mostly lost overnight. This pattern of change seems to be characteristic of development, as it is absent in adults (Extended Data Figs. 5, 6).

Movement along the DiSC also occurs on timescales that are faster than hours, namely within bouts of singing—that is, groups of vocalizations that are preceded and followed by a pause (average bout duration  $3.81 \pm 0.83$  s across birds). We subdivide each bout into ten consecutive periods, compute pooled neighbourhood times for each period (over all bouts in a day), and track change through the corresponding percentiles (Fig. 3c, d). Within bouts, large changes along the DiSC occur at



**Fig. 3 | Multiple components of behavioural change during sensory-motor learning.** **a**, Average repertoire dating percentiles (for five birds) describing within and across-day changes along the DiSC. For each production day and period, five percentiles of the pooled neighbourhood times (Fig. 2c) are arranged vertically (lines). **b**, Average of data from **a** across days 60–70, expressed relative to the average 50th percentile. **c**, Within-bout changes. As for **a**, but based on production day and period in a singing bout. **d**, As for **b**, but averaged across data from **c**. **e**, Span and shift for the 5th, 50th and 95th percentiles (blue arrows in **b**, analogous to Fig. 1c) averaged over days 50–80, separately for syllables (points) and birds (colours). Black lines indicate medians and 95% bootstrapped confidence intervals over all points. **h**, Regressions. **i**, Typical. **j**, Anticipations. **k**, Direction of slow change. **l**, Within-day change. **m**, Across-day change. **n**, Slow 1. **o**, Slow 2. **p**, Typical. **q**, Anticipations. **r**, Within-day change. **s**, Across-day change. **t**, Slow 1. **u**, Slow 2. **v**, Typical. **w**, Anticipations. **x**, Within-day change. **y**, Across-day change. **z**, Slow 1. **aa**, Slow 2. **ab**, Typical. **ac**, Anticipations. **ad**, Within-day change. **ae**, Across-day change. **af**, Slow 1. **ag**, Slow 2. **ah**, Typical. **ai**, Anticipations. **aj**, Within-day change. **ak**, Across-day change. **al**, Slow 1. **am**, Slow 2. **an**, Typical. **ao**, Anticipations. **ap**, Within-day change. **aq**, Across-day change. **ar**, Slow 1. **as**, Slow 2. **at**, Typical. **au**, Anticipations. **av**, Within-day change. **aw**, Across-day change. **ax**, Slow 1. **ay**, Slow 2. **az**, Typical. **ba**, Anticipations. **bb**, Within-day change. **bc**, Across-day change. **bd**, Slow 1. **be**, Slow 2. **bf**, Typical. **bg**, Anticipations. **bh**, Within-day change. **bi**, Across-day change. **bj**, Slow 1. **bk**, Slow 2. **bl**, Typical. **bm**, Anticipations. **bn**, Within-day change. **bo**, Across-day change. **bp**, Slow 1. **bq**, Slow 2. **br**, Typical. **bs**, Anticipations. **bt**, Within-day change. **bu**, Across-day change. **bv**, Slow 1. **bv**, Slow 2. **bw**, Typical. **bx**, Anticipations. **by**, Within-day change. **bz**, Across-day change. **ca**, Slow 1. **cb**, Slow 2. **cc**, Typical. **cd**, Anticipations. **ce**, Within-day change. **cf**, Across-day change. **cg**, Slow 1. **ch**, Slow 2. **ci**, Typical. **cj**, Anticipations. **ck**, Within-day change. **cl**, Across-day change. **cm**, Slow 1. **cn**, Slow 2. **co**, Typical. **cp**, Anticipations. **cq**, Within-day change. **cr**, Across-day change. **cs**, Slow 1. **ct**, Slow 2. **cu**, Typical. **cv**, Anticipations. **cw**, Within-day change. **cx**, Across-day change. **cy**, Slow 1. **cz**, Slow 2. **ca**, Typical. **cb**, Anticipations. **cc**, Within-day change. **cd**, Across-day change. **ce**, Slow 1. **cf**, Slow 2. **cg**, Typical. **ch**, Anticipations. **ci**, Within-day change. **cj**, Across-day change. **ck**, Slow 1. **cl**, Slow 2. **cm**, Typical. **cn**, Anticipations. **co**, Within-day change. **cp**, Across-day change. **cq**, Slow 1. **cr**, Slow 2. **cs**, Typical. **ct**, Anticipations. **cu**, Within-day change. **cv**, Across-day change. **cw**, Slow 1. **cx**, Slow 2. **cy**, Typical. **cz**, Anticipations. **da**, Within-day change. **db**, Across-day change. **dc**, Slow 1. **dd**, Slow 2. **de**, Typical. **df**, Anticipations. **dg**, Within-day change. **dh**, Across-day change. **di**, Slow 1. **dj**, Slow 2. **dk**, Typical. **dl**, Anticipations. **dm**, Within-day change. **dn**, Across-day change. **do**, Slow 1. **dp**, Slow 2. **dq**, Typical. **dr**, Anticipations. **ds**, Within-day change. **dt**, Across-day change. **du**, Slow 1. **dv**, Slow 2. **dw**, Typical. **dx**, Anticipations. **dy**, Within-day change. **dz**, Across-day change. **ea**, Slow 1. **eb**, Slow 2. **ec**, Typical. **ed**, Anticipations. **ee**, Within-day change. **ef**, Across-day change. **eg**, Slow 1. **eh**, Slow 2. **ei**, Typical. **ej**, Anticipations. **ek**, Within-day change. **el**, Across-day change. **em**, Slow 1. **en**, Slow 2. **eo**, Typical. **ep**, Anticipations. **eq**, Within-day change. **er**, Across-day change. **es**, Slow 1. **et**, Slow 2. **eu**, Typical. **ev**, Anticipations. **ew**, Within-day change. **ex**, Across-day change. **ey**, Slow 1. **ez**, Slow 2. **fa**, Typical. **fb**, Anticipations. **fc**, Within-day change. **fd**, Across-day change. **fe**, Slow 1. **ff**, Slow 2. **fg**, Typical. **fh**, Anticipations. **fi**, Within-day change. **fj**, Across-day change. **fk**, Slow 1. **fl**, Slow 2. **fm**, Typical. **fn**, Anticipations. **fo**, Within-day change. **fp**, Across-day change. **fq**, Slow 1. **fr**, Slow 2. **fs**, Typical. **ft**, Anticipations. **fu**, Within-day change. **fv**, Across-day change. **fw**, Slow 1. **fx**, Slow 2. **fy**, Typical. **fz**, Anticipations. **ga**, Anticipations. **gb**, Within-day change. **gc**, Across-day change. **gd**, Slow 1. **ge**, Slow 2. **gf**, Typical. **gg**, Anticipations. **gh**, Within-day change. **gi**, Across-day change. **gj**, Slow 1. **gk**, Slow 2. **gl**, Typical. **gm**, Anticipations. **gn**, Within-day change. **go**, Across-day change. **gp**, Slow 1. **gq**, Slow 2. **gr**, Typical. **gs**, Anticipations. **gt**, Within-day change. **gu**, Across-day change. **gv**, Slow 1. **gw**, Slow 2. **gx**, Typical. **gy**, Anticipations. **gz**, Within-day change. **ha**, Across-day change. **hb**, Slow 1. **hc**, Slow 2. **hd**, Typical. **he**, Anticipations. **hf**, Within-day change. **hg**, Across-day change. **hh**, Slow 1. **hi**, Slow 2. **hj**, Typical. **hk**, Anticipations. **hl**, Within-day change. **hm**, Across-day change. **hn**, Slow 1. **ho**, Slow 2. **hp**, Typical. **hq**, Anticipations. **hr**, Within-day change. **hs**, Across-day change. **ht**, Slow 1. **hu**, Slow 2. **hv**, Typical. **hw**, Anticipations. **hx**, Within-day change. **hy**, Across-day change. **hz**, Slow 1. **ia**, Slow 2. **ib**, Typical. **ic**, Anticipations. **id**, Within-day change. **ie**, Across-day change. **if**, Slow 1. **ig**, Slow 2. **ih**, Typical. **ii**, Anticipations. **ij**, Within-day change. **ik**, Across-day change. **il**, Slow 1. **im**, Slow 2. **in**, Typical. **io**, Anticipations. **ip**, Within-day change. **iq**, Across-day change. **ir**, Slow 1. **is**, Slow 2. **it**, Typical. **iu**, Anticipations. **iv**, Within-day change. **iw**, Across-day change. **ix**, Slow 1. **iy**, Slow 2. **iz**, Typical. **ja**, Anticipations. **jb**, Within-day change. **jc**, Across-day change. **jd**, Slow 1. **je**, Slow 2. **jf**, Typical. **jg**, Anticipations. **jh**, Within-day change. **ji**, Across-day change. **jj**, Slow 1. **jk**, Slow 2. **jl**, Typical. **jm**, Anticipations. **jn**, Within-day change. **jo**, Across-day change. **jp**, Slow 1. **jq**, Slow 2. **jr**, Typical. **js**, Anticipations. **jt**, Within-day change. **ju**, Across-day change. **jv**, Slow 1. **jw**, Slow 2. **jx**, Typical. **gy**, Anticipations. **gz**, Within-day change. **ka**, Across-day change. **kb**, Slow 1. **kc**, Slow 2. **kd**, Typical. **ke**, Anticipations. **kf**, Within-day change. **kg**, Across-day change. **kh**, Slow 1. **ki**, Slow 2. **kl**, Typical. **km**, Anticipations. **kn**, Within-day change. **ko**, Across-day change. **kp**, Slow 1. **kq**, Slow 2. **kr**, Typical. **ks**, Anticipations. **kt**, Within-day change. **ku**, Across-day change. **kv**, Slow 1. **kw**, Slow 2. **kx**, Typical. **ky**, Anticipations. **kz**, Within-day change. **la**, Across-day change. **lb**, Slow 1. **lc**, Slow 2. **ld**, Typical. **le**, Anticipations. **lf**, Within-day change. **lg**, Across-day change. **lh**, Slow 1. **li**, Slow 2. **lj**, Typical. **lk**, Anticipations. **ll**, Within-day change. **lm**, Across-day change. **ln**, Slow 1. **lo**, Slow 2. **lp**, Typical. **lq**, Anticipations. **lr**, Within-day change. **ls**, Across-day change. **lt**, Slow 1. **lu**, Slow 2. **lv**, Typical. **lw**, Anticipations. **lx**, Within-day change. **ly**, Across-day change. **lz**, Slow 1. **ma**, Slow 2. **mb**, Typical. **mc**, Anticipations. **md**, Within-day change. **me**, Across-day change. **mf**, Slow 1. **mg**, Slow 2. **mh**, Typical. **mi**, Anticipations. **mj**, Within-day change. **mk**, Across-day change. **ml**, Slow 1. **mn**, Slow 2. **mo**, Typical. **mp**, Anticipations. **mq**, Within-day change. **mr**, Across-day change. **ms**, Slow 1. **mt**, Slow 2. **mu**, Typical. **mv**, Anticipations. **mw**, Within-day change. **mx**, Across-day change. **my**, Slow 1. **mz**, Slow 2. **na**, Typical. **nb**, Anticipations. **nc**, Within-day change. **nd**, Across-day change. **ne**, Slow 1. **nf**, Slow 2. **ng**, Typical. **nh**, Anticipations. **ni**, Within-day change. **nj**, Across-day change. **nk**, Slow 1. **nl**, Slow 2. **nm**, Typical. **no**, Anticipations. **np**, Within-day change. **nq**, Across-day change. **nr**, Slow 1. **ns**, Slow 2. **nt**, Typical. **nu**, Anticipations. **nv**, Within-day change. **nw**, Across-day change. **nx**, Slow 1. **ny**, Slow 2. **nz**, Typical. **oa**, Anticipations. **ob**, Within-day change. **oc**, Across-day change. **od**, Slow 1. **oe**, Slow 2. **of**, Typical. **og**, Anticipations. **oh**, Within-day change. **oi**, Across-day change. **oj**, Slow 1. **ok**, Slow 2. **ol**, Typical. **om**, Anticipations. **on**, Within-day change. **oo**, Across-day change. **op**, Slow 1. **oq**, Slow 2. **or**, Typical. **os**, Anticipations. **ot**, Within-day change. **ou**, Across-day change. **ov**, Slow 1. **ow**, Slow 2. **ox**, Typical. **oy**, Anticipations. **oz**, Within-day change. **pa**, Across-day change. **pb**, Slow 1. **pc**, Slow 2. **pd**, Typical. **pe**, Anticipations. **pf**, Within-day change. **pg**, Across-day change. **ph**, Slow 1. **pi**, Slow 2. **pj**, Typical. **pk**, Anticipations. **pl**, Within-day change. **pm**, Across-day change. **pn**, Slow 1. **po**, Slow 2. **pp**, Typical. **pq**, Anticipations. **pr**, Within-day change. **ps**, Across-day change. **pt**, Slow 1. **pu**, Slow 2. **pv**, Typical. **pw**, Anticipations. **px**, Within-day change. **py**, Across-day change. **pz**, Slow 1. **qa**, Slow 2. **qb**, Typical. **qc**, Anticipations. **qd**, Within-day change. **qe**, Across-day change. **qf**, Slow 1. **qg**, Slow 2. **qh**, Typical. **qi**, Anticipations. **qj**, Within-day change. **qk**, Across-day change. **ql**, Slow 1. **qm**, Slow 2. **qn**, Typical. **qo**, Anticipations. **qp**, Within-day change. **qq**, Across-day change. **qr**, Slow 1. **qs**, Slow 2. **qt**, Typical. **qu**, Anticipations. **qv**, Within-day change. **qw**, Across-day change. **qx**, Slow 1. **qy**, Slow 2. **qz**, Typical. **ra**, Anticipations. **rb**, Within-day change. **rc**, Across-day change. **rd**, Slow 1. **re**, Slow 2. **rf**, Typical. **rg**, Anticipations. **rh**, Within-day change. **ri**, Across-day change. **rj**, Slow 1. **rk**, Slow 2. **rl**, Typical. **rm**, Anticipations. **rn**, Within-day change. **ro**, Across-day change. **rp**, Slow 1. **rq**, Slow 2. **rr**, Typical. **rs**, Anticipations. **rt**, Within-day change. **ru**, Across-day change. **rv**, Slow 1. **rw**, Slow 2. **rx**, Typical. **ry**, Anticipations. **rz**, Within-day change. **sa**, Across-day change. **sb**, Slow 1. **sc**, Slow 2. **sd**, Typical. **se**, Anticipations. **sf**, Within-day change. **sg**, Across-day change. **sh**, Slow 1. **si**, Slow 2. **sj**, Typical. **sk**, Anticipations. **sl**, Within-day change. **sm**, Across-day change. **sn**, Slow 1. **so**, Slow 2. **sp**, Typical. **sq**, Anticipations. **sr**, Within-day change. **ss**, Across-day change. **st**, Slow 1. **su**, Slow 2. **sv**, Typical. **sw**, Anticipations. **sx**, Within-day change. **sy**, Across-day change. **sz**, Slow 1. **ta**, Slow 2. **tb**, Typical. **tc**, Anticipations. **td**, Within-day change. **te**, Across-day change. **tf**, Slow 1. **tg**, Slow 2. **th**, Typical. **ti**, Anticipations. **tj**, Within-day change. **tk**, Across-day change. **tl**, Slow 1. **tm**, Slow 2. **tn**, Typical. **to**, Anticipations. **tp**, Within-day change. **tp**, Across-day change. **tr**, Slow 1. **ts**, Slow 2. **tt**, Typical. **tu**, Anticipations. **tv**, Within-day change. **tv**, Across-day change. **tw**, Slow 1. **tw**, Slow 2. **tx**, Typical. **ty**, Anticipations. **tz**, Within-day change. **ua**, Across-day change. **ub**, Slow 1. **uc**, Slow 2. **ud**, Typical. **ue**, Anticipations. **uf**, Within-day change. **uf**, Across-day change. **ug**, Slow 1. **ug**, Slow 2. **uh**, Typical. **ui**, Anticipations. **uj**, Within-day change. **uj**, Across-day change. **uk**, Slow 1. **uk**, Slow 2. **ul**, Typical. **um**, Anticipations. **un**, Within-day change. **un**, Across-day change. **uo**, Slow 1. **uo**, Slow 2. **up**, Typical. **uq**, Anticipations. **ur**, Within-day change. **ur**, Across-day change. **us**, Slow 1. **us**, Slow 2. **ut**, Typical. **uv**, Anticipations. **uv**, Within-day change. **uv**, Across-day change. **uw**, Slow 1. **uw**, Slow 2. **ux**, Typical. **uy**, Anticipations. **uz**, Within-day change. **va**, Across-day change. **vb**, Slow 1. **vc**, Slow 2. **vd**, Typical. **ve**, Anticipations. **vf**, Within-day change. **vf**, Across-day change. **vg**, Slow 1. **vg**, Slow 2. **vh**, Typical. **vi**, Anticipations. **vi**, Within-day change. **vi**, Across-day change. **vj**, Slow 1. **vj**, Slow 2. **vk**, Typical. **vl**, Anticipations. **vl**, Within-day change. **vl**, Across-day change. **vm**, Slow 1. **vm**, Slow 2. **vn**, Typical. **vo**, Anticipations. **vo**, Within-day change. **vo**, Across-day change. **vp**, Slow 1. **vp**, Slow 2. **vq**, Typical. **vr**, Anticipations. **vr**, Within-day change. **vr**, Across-day change. **vs**, Slow 1. **vs**, Slow 2. **vt**, Typical. **vu**, Anticipations. **vu**, Within-day change. **vu**, Across-day change. **vv**, Slow 1. **vv**, Slow 2. **vw**, Typical. **vx**, Anticipations. **vx**, Within-day change. **vx**, Across-day change. **vy**, Slow 1. **vy**, Slow 2. **vz**, Typical. **wa**, Anticipations. **wb**, Within-day change. **wc**, Across-day change. **wd**, Slow 1. **we**, Slow 2. **wf**, Typical. **wg**, Anticipations. **wh**, Within-day change. **wh**, Across-day change. **wi**, Slow 1. **wi**, Slow 2. **wj**, Typical. **wk**, Anticipations. **wl**, Within-day change. **wl**, Across-day change. **wm**, Slow 1. **wm**, Slow 2. **wn**, Typical. **wo**, Anticipations. **wo**, Within-day change. **wo**, Across-day change. **wp**, Slow 1. **wp**, Slow 2. **wq**, Typical. **wr**, Anticipations. **wr**, Within-day change. **wr**, Across-day change. **ws**, Slow 1. **ws**, Slow 2. **wt**, Typical. **wu**, Anticipations. **wu**, Within-day change. **wu**, Across-day change. **wv**, Slow 1. **wv**, Slow 2. **wx**, Typical. **wy**, Anticipations. **wz**, Within-day change. **xa**, Across-day change. **xb**, Slow 1. **xc**, Slow 2. **xd**, Typical. **xe**, Anticipations. **xf**, Within-day change. **xf**, Across-day change. **yg**, Slow 1. **yg**, Slow 2. **yh**, Typical. **yi**, Anticipations. **yi**, Within-day change. **yi**, Across-day change. **yj**, Slow 1. **yj**, Slow 2. **yk**, Typical. **yl**, Anticipations. **yl**, Within-day change. **yl**, Across-day change. **ym**, Slow 1. **ym**, Slow 2. **yn**, Typical. **yo**, Anticipations. **yo**, Within-day change. **yo**, Across-day change. **yp**, Slow 1. **yp**, Slow 2. **yq**, Typical. **yr**, Anticipations. **yr**, Within-day change. **yr**, Across-day change. **ys**, Slow 1. **ys**, Slow 2. **yt**, Typical. **yu**, Anticipations. **yu**, Within-day change. **yu**, Across-day change. **yv**, Slow 1. **yv**, Slow 2. **yz**, Typical. **za**, Anticipations. **zb**, Within-day change. **zc**, Across-day change. **zd**, Slow 1. **ze**, Slow 2. **zf**, Typical. **zg**, Anticipations. **zh**, Within-day change. **zh**, Across-day change. **zi**, Slow 1. **zi**, Slow 2. **zj**, Typical. **zk**, Anticipations. **zk**, Within-day change. **zk**, Across-day change. **zl**, Slow 1. **zl**, Slow 2. **zm**, Typical. **zn**, Anticipations. **zn**, Within-day change. **zn**, Across-day change. **zo**, Slow 1. **zo**, Slow 2. **zp**, Typical. **zq**, Anticipations. **zr**, Within-day change. **zr**, Across-day change. **zs**, Slow 1. **zs**, Slow 2. **zt**, Typical. **zu**, Anticipations. **zu**, Within-day change. **zu**, Across-day change. **zv**, Slow 1. **zv**, Slow 2. **zz**, Typical.

**f**, Simulated stratified mixing matrices (right) for three models (left) of the alignment of within-day and across-day change with the DiSC. **g**, Average measured stratified mixing matrices (five birds, days 60–70). **h–j**, Stratified behavioural trajectory based on **g**. Different two-dimensional projections reveal the DiSC (**h**), as well as within-day (**i**) and across-day (**j**) change not aligned with the DiSC (labels 1–5 represent different strata). The full ten-dimensional trajectories faithfully reproduce the structure of the stratified mixing matrices (MDS stress = 0.016); the depicted four-dimensional subspace captures 81% of the ten-dimensional variance. **k**, Separate projections for each stratum onto the local DiSC (black arrows in upper diagrams; points represent strata from **h**).

the regressive tail of the behavioural repertoire: vocalizations are most regressive at the onset and offset of bouts (Fig. 3c, d, 5th percentile). Similar, albeit weaker changes occur for typical renditions (Fig. 3c, d, red). The same apparent changes in song maturity are observed when short and long bouts (durations  $2.30 \pm 0.54$  s versus  $6.28 \pm 1.73$  s) are considered separately. Song maturity thus decreases at the end of a bout, not after a fixed time into the bout (Extended Data Fig. 5a–c).

### Misaligned behavioural components

The repertoire time reveals within-day and within-bout changes that mirror, on a faster timescale, changes that also occur over many days (see Supplementary Methods). As above (Fig. 1), we refer to such components of change as being aligned with the DiSC, and to components that are not reflected in the repertoire time as being misaligned.

We identify both aligned and misaligned components of change through the ‘stratified mixing matrix’, which combines a neighbourhood-mixing matrix (for example, Fig. 2f) with repertoire dating. Each day’s behavioural repertoire is binned into five consecutive production periods. Within each period, the behavioural repertoire is subdivided into five strata on the basis of repertoire time (Fig. 2d, quintiles). All renditions from a day thus fall into  $5 \times 5 = 25$  bins. The stratified mixing matrix measures similarity between 50 bins that combine the data from two adjacent days (Fig. 3g). We compare the measured stratified mixing matrix with simulations that differ with respect to how within-day change and change across adjacent days align with the DiSC (Fig. 3f and Extended Data Fig. 4e–j). In model 1, development is one-dimensional and therefore aligned with the DiSC (Fig. 3f, top; similar to Fig. 1e). In model 2, within-day change involves a component that is not aligned with the DiSC (Fig. 3f, middle; similar to Fig. 1g). In model 3, adjacent days are separated not only along the DiSC, but also along a direction

orthogonal to both the DiSC and the direction of within-day change (Fig. 3f, bottom, across-day change). Prominent ‘stripes’ along every other diagonal in the measured mixing matrix (Fig. 3g) indicate a larger similarity between renditions from the same day than between renditions from adjacent days, as predicted by model 3, suggesting that several misaligned components contribute to change at fast timescales.

From the stratified mixing matrix, we infer stratified behavioural trajectories. The two-dimensional projection that captures most of the variance due to strata (Fig. 3h) resembles Fig. 2h and reflects the DiSC. Consistent with repertoire dating, behavioural change along the DiSC between adjacent days (Fig. 3h, blue versus red for each stratum) is small compared with the spread of the behaviour for one day along the DiSC (for example, blue points, strata 1–5). For each stratum, however, much of the change that occurs within a day is misaligned with the DiSC (Fig. 3i, k; early versus late separated along the orthogonal dimension of within-day change). Yet another misaligned component is necessary to appropriately capture change across adjacent days (Fig. 3j). These properties of aligned and misaligned components are replicated by a linear analysis based on spectral features that are chosen to capture change at specific timescales (Extended Data Figs. 7, 8), and are robust to how song is parameterized and segmented, and to how nearest neighbours are defined (Extended Data Figs. 9, 10).

### Discussion

Our analysis of high-dimensional vocalizations reveals that vocal learning and development do not reflect an underlying one-dimensional process. Single behavioural features in isolation therefore provide an incomplete account of behavioural change during development and learning. The weak consolidation observed here (Fig. 1d) and elsewhere<sup>2,14</sup> at the level of single features appears to reflect prominent



misaligned components of within-day change, rather than weak consolidation along the DiSC (Fig. 1h, i). Strong overnight consolidation along the DiSC across much of the behavioural repertoire (Fig. 3a, b) seems consistent with consolidation patterns observed for skilled motor learning in humans<sup>23,24,28</sup> and of motor adaptation in humans<sup>1,18</sup> and birds<sup>4</sup>.

Our characterization of behaviour on the basis of nearest-neighbour statistics can be applied when no accurate parametric model of the behaviour is known, as is the case at present for most natural, complex behaviours. The approach is largely complementary to methods that rely on clustering behaviour into distinct categories<sup>2,3,10,29</sup>. Forgoing an explicit clustering of the data can be advantageous, because assuming the existence of clusters can be an unwarranted approximation<sup>30</sup> and may impede the characterization of behaviour that appears not to be clustered (such as juvenile zebra finch song; Extended Data Fig. 1); moreover, determining correct cluster boundaries is in general an ill-defined problem<sup>30</sup>. Notably, our analyses require only an indicator function that selects nearest neighbours (based here on a ‘locally meaningful’ distance metric)—a much weaker requirement than a globally valid distance metric or the existence of a low-dimensional feature space that globally maps behavioural space<sup>11</sup>. These properties make repertoire dating applicable to almost any behaviour and other high-dimensional datasets, including data that are characterized by ‘labels’ other than production time. Repertoire dating may thus provide a general account of learning and change that is amenable to comparisons between different behaviours and model systems, including different species<sup>17</sup> and artificial systems<sup>5</sup>.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-019-1892-x>.

1. Brashers-Krug, T., Shadmehr, R. & Bizzi, E. Consolidation in human motor memory. *Nature* **382**, 252–255 (1996).
2. Derégnaucourt, S., Mitra, P. P., Fehér, O., Pytte, C. & Tchernichovski, O. How sleep affects the developmental learning of bird song. *Nature* **433**, 710–716 (2005).
3. Tchernichovski, O., Mitra, P. P., Lints, T. & Nottebohm, F. Dynamics of the vocal imitation process: how a zebra finch learns its song. *Science* **291**, 2564–2569 (2001).
4. Andalman, A. S. & Fee, M. S. A basal ganglia–forebrain circuit in the songbird biases motor output to avoid vocal errors. *Proc. Natl Acad. Sci. USA* **106**, 12518–12523 (2009).
5. Arulkumaran, K., Deisenroth, M. P., Brundage, M. & Bharath, A. A. Deep reinforcement learning: a brief survey. *IEEE Signal Process. Mag.* **34**, 26–38 (2017).

6. Ingram, J. N., Flanagan, J. R. & Wolpert, D. M. Context-dependent decay of motor memories during skill acquisition. *Curr. Biol.* **23**, 1107–1112 (2013).
7. Klaus, A. et al. The spatiotemporal organization of the striatum encodes action space. *Neuron* **95**, 1171–1180 (2017).
8. Han, S., Taralova, E., Dupre, C. & Yuste, R. Comprehensive machine learning analysis of *Hydra* behavior reveals a stable basal behavioral repertoire. *eLife* **7**, e32605 (2018).
9. Egnor, S. E. R. & Branson, K. Computational analysis of behavior. *Annu. Rev. Neurosci.* **39**, 217–236 (2016).
10. Wiltischko, A. B. et al. Mapping sub-second structure in mouse behavior. *Neuron* **88**, 1121–1135 (2015).
11. van der Maaten, L. Accelerating t-SNE using tree-based algorithms. *J. Mach. Learn. Res.* **15**, 3221–3245 (2014).
12. Chen, H. & Friedman, J. J. H. A new graph-based two-sample test for multivariate and object data. *J. Am. Stat. Assoc.* **145**, 1–41 (2016).
13. Hawks, M. *Graph-Theoretic Statistical Methods for Detecting and Localizing Distributional Change in Multivariate Data*. PhD Thesis, Naval Postgraduate School, Monterey, California (2015).
14. Shank, S. S. & Margoliash, D. Sleep and sensorimotor integration during early vocal learning in a songbird. *Nature* **458**, 73–77 (2009).
15. Fenn, K. M., Nusbaum, H. C. & Margoliash, D. Consolidation during sleep of perceptual learning of spoken language. *Nature* **425**, 614–616 (2003).
16. Tchernichovski, O., Nottebohm, F., Ho, C. E., Pesaran, B. & Mitra, P. P. A procedure for an automated measurement of song similarity. *Anim. Behav.* **59**, 1167–1176 (2000).
17. Anderson, D. J. J. & Perona, P. Toward a science of computational ethology. *Neuron* **84**, 18–31 (2014).
18. Krakauer, J. W. & Shadmehr, R. Consolidation of motor memory. *Trends Neurosci.* **29**, 58–64 (2006).
19. Brainard, M. S. & Doupe, A. J. What songbirds teach us about learning. *Nature* **417**, 351–358 (2002).
20. Catchpole, C. K. & Slater, P. J. B. *Bird Song: Biological Themes and Variations* (Cambridge Univ. Press, 2003).
21. Dhawale, A. K., Smith, M. A. & Ölveczky, B. P. The role of variability in motor learning. *Annu. Rev. Neurosci.* **40**, 479–498 (2017).
22. Lipkind, D. et al. Songbirds work around computational complexity by learning song vocabulary independently of sequence. *Nat. Commun.* **8**, 1247 (2017).
23. Korman, M. et al. Daytime sleep condenses the time course of motor memory consolidation. *Nat. Neurosci.* **10**, 1206–1213 (2007).
24. Fischer, S., Hallschmid, M., Elsner, A. L. & Born, J. Sleep forms memory for finger skills. *Proc. Natl Acad. Sci. USA* **99**, 11987–11991 (2002).
25. Kruskal, J. B. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika* **29**, 1–27 (1964).
26. Fehér, O., Wang, H., Saar, S., Mitra, P. P. & Tchernichovski, O. De novo establishment of wild-type song culture in the zebra finch. *Nature* **459**, 564–568 (2009).
27. Adam, I. & Elemans, C. P. H. Vocal motor performance in birdsong requires brain–body interaction. *eNeuro* **6**, ENEURO.0053-19.2019 (2019).
28. Walker, M. P., Brakefield, T., Hobson, J. A. & Stickgold, R. Dissociable stages of human memory consolidation and reconsolidation. *Nature* **425**, 616–620 (2003).
29. Vogelstein, J. T. et al. Discovery of brainwide neural-behavioral maps via multiscale unsupervised structure learning. *Science* **344**, 386–392 (2014).
30. Fahad, A. et al. A survey of clustering algorithms for big data: taxonomy and empirical analysis. *IEEE Trans. Emerg. Top. Comput.* **2**, 267–279 (2014).

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2020

## Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

## Data availability

The data that support the findings of this study are available from the corresponding authors upon reasonable request.

## Code availability

We provide source code for our nearest-neighbour-based analyses (<https://github.com/skollmor/repertoireDating>). We also provide a data-browser/data analysis graphical user interface that we used to perform some analyses and to explore and visualize data (<https://github.com/skollmor/dspace>).

**Acknowledgements** We thank J. Herbst and Z. Huang for performing the tutoring experiments; and A. Zai, S. Surace, A. Huber, I. Calangiu and K. Martin for discussing the manuscript. This work was supported by grants from the Simons Foundation (SCGB 328189 and SCGB 543013 to V.M.) and the Swiss National Science Foundation (SNSF PP00P3\_157539 to V.M. and SNSF 31003A\_182638 to R. H.). All experimental procedures were approved by the Veterinary Office of the Canton of Zurich.

**Author contributions** S.K. conceived of the approach. S.K. and V.M. performed analyses. S.K., V.M. and R.H.R.H. wrote the paper. R.H.R.H. conceived and supervised collection of the behavioural data.

**Competing interests** The authors declare no competing interests.

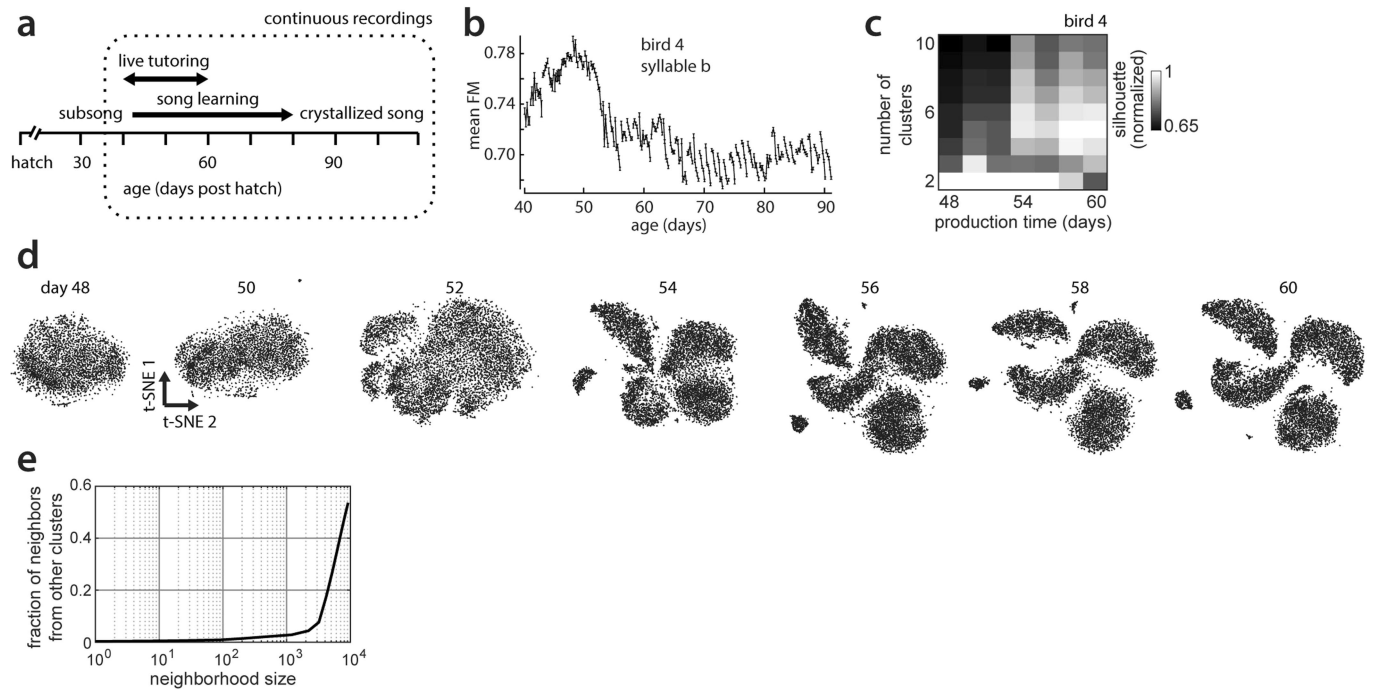
## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41586-019-1892-x>.

**Correspondence and requests for materials** should be addressed to S.K. or V.M.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>.

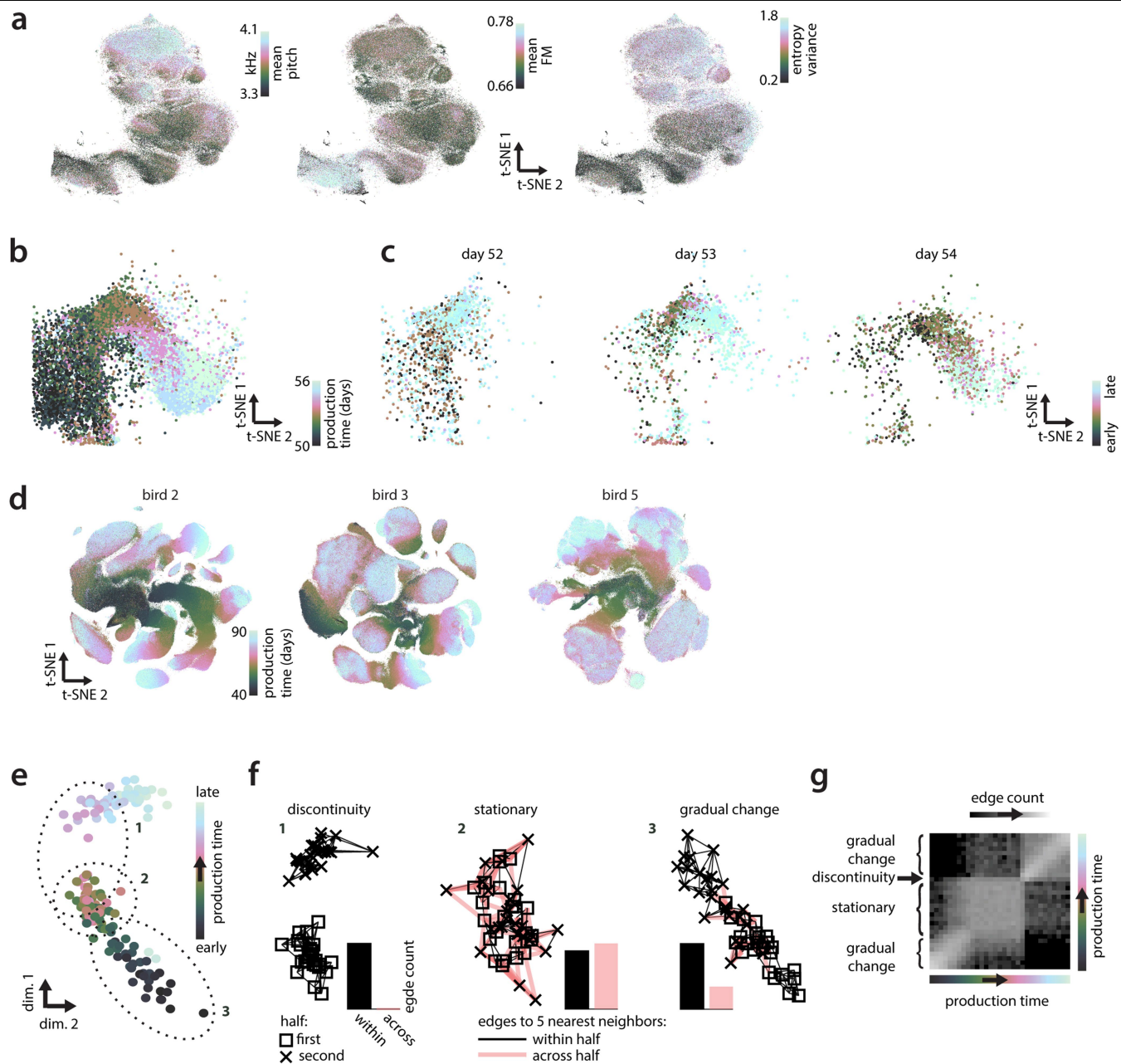




### Extended Data Fig. 1 | Clustering of juvenile and adult zebra finch song.

**a**, Vocal development in male zebra finches. Tutoring by an adult male started at around day 46 (post-hatch) and lasted 10–20 days. **b**, Time course of the acoustic feature frequency modulation (FM), for syllable b in the example bird (compare with Fig. 1b). **c**, Normalized mean silhouette values for 2–10 clusters for vocalizations from the seven days shown in **d**. High values indicate evidence for the respective cluster count. Normalized mean silhouette coefficients are based on 20 repetitions of *k*-means clustering of random subsets of 1,000 68-ms onset-aligned spectrogram segments from a single day (as in **d**), projected onto the first five principal components. **d**, *t*-SNE visualizations of vocalizations produced on a given day post-hatch for the example bird (bird 4, the same bird as in Fig. 1a–c). A separate embedding was computed for each day, and the embedding's initial condition was based on the previous day. Note

the gradual emergence of clusters, each corresponding to a distinct syllable type (for example, syllables i, a, b, c in Fig. 2a). **e**, Average fraction of neighbours from a different cluster, as a function of neighbourhood size. These data are analogous to those from **c**, **d** but for vocalizations from day 90 (12,854 data points), when clusters are fully developed. For a wide range of neighbourhood sizes, the neighbours of a data point mostly belong to the same cluster or syllable type. For a neighbourhood size of 100, the average fraction of out-of-cluster neighbours from the same day is 0.0089. Thus, for an appropriately chosen neighbourhood size, nearest-neighbour methods respect clustering structure in the data by construction, and sidestep having to explicitly identify clusters in the data. In most analyses, we computed nearest neighbours for data from all days, meaning that clustering structure is respected even for neighbourhood sizes that are slightly larger than those suggested in **e**.

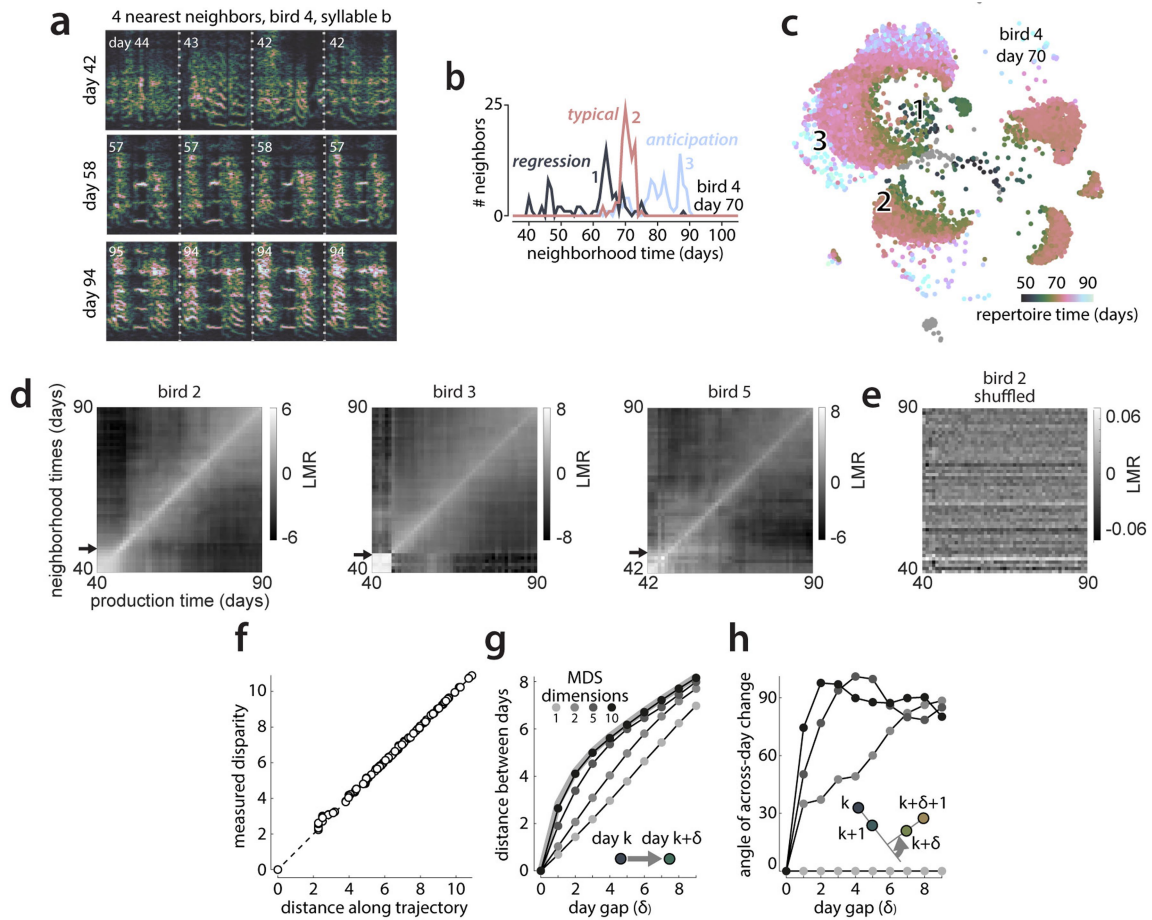


**Extended Data Fig. 2** | See next page for caption.

# Article

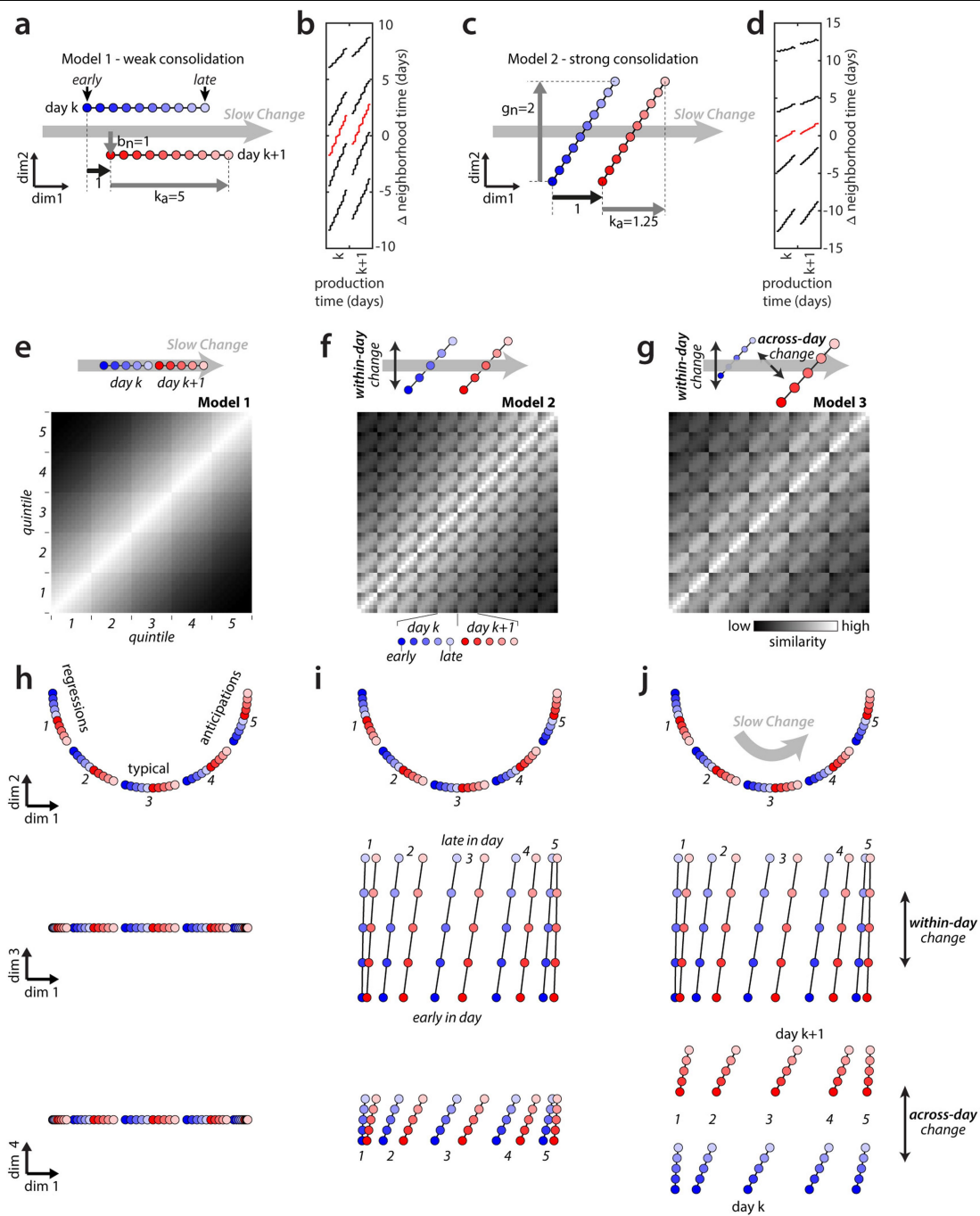
**Extended Data Fig. 2 | Properties of large-scale embeddings.** **a**, Three auditory features computed on renditions of syllable **b**. This panel uses the same embedding as in Fig. 2a, but with different colours. **b**, Across-day change in vocalizations. This is a magnified cutout from the bottom left region of the dashed outline of Fig. 2a. The colours differ from Fig. 2a, and points from days 50–56 only are shown. **c**, Within-day change in vocalizations. Points from **b** are shown separately for three individual days and coloured according to production time within the day (early to late). Vocalizations change within a day: early vocalizations (dark green) are more similar to vocalizations from previous days (dark green points in **b**); late vocalizations (light blue) are more similar to vocalizations from future days (light blue in **b**). **d**, *t*-SNE visualizations for dense recordings from three birds (analogous to Fig. 2a). **e–g**, Illustration of a fictitious behaviour that undergoes distinct phases of abrupt change, no change and gradual change, and the identification of these phases on the basis of nearest-neighbour graphs. **e**, A low-dimensional representation of the behaviour. Each point corresponds to a behavioural rendition (for example, a syllable rendition) and is coloured according to production time. Similar renditions (for example, syllable renditions with similar spectrograms) appear near each other in this representation. The dotted ellipses mark three subsets of points corresponding to: (1) a phase of abrupt change; (2) a phase of no change; and (3) a phase of gradual change. **f**, Nearest-neighbour graphs for the three subsets of points in **e**. Points are replotted from **e** with different symbols,

indicating whether their production times fall within the first half (squares) or second half (crosses) of the corresponding subset. Edges connect each point to its five nearest neighbours. The edge colour marks neighbouring pairs of points falling into the same (black) or different (red) halves. Relative counts of within- and across-half edges differ according to the nature of the underlying behavioural change (histograms of edge counts). If an abrupt change in behaviour occurs between the first and second half, nearest neighbours of points in one half will all be points from the same half, and none from the other half (discontinuity). When behaviour is stationary, the neighbourhoods are maximally mixed: that is, every point has about an equal number of neighbours from the two halves. Phases of gradual change result in intermediate levels of mixing. **g**, Mixing matrix for the simulated data in **e**, analogous to Fig. 2e. Each location in the matrix corresponds to a pair of production times. Strong mixing (white) indicates a large number of nearest-neighbour edges across the two corresponding production times (as in **f**; stationary) and thus similar behaviour at the two times. Weak mixing (black) indicates a small number of such edges (as in **f**; discontinuity), and thus dissimilar behaviour. Note that such statistics on the composition of local neighbourhoods can be computed for any kind of behaviour and are invariant with respect to transformations of the data that preserve nearest neighbours, such as scaling, translation and rotation. These properties make nearest-neighbour approaches highly general.



**Extended Data Fig. 3 | Repertoire dating and the direction of slow change.** **a**, The four nearest neighbours for example vocalizations (bird 4, syllable b, from Fig. 1a). Production times of nearest neighbours (numbers) need not equal that of the corresponding example rendition. **b**, Neighbourhood production times for three renditions from day 70 (analogous to Fig. 2a, inset). Rendition 2 is 'typical' for day 70 (most neighbours lie in the same or adjacent days); renditions 1 and 3 are a 'regression' and an 'anticipation' (with neighbours predominantly produced in the past or future, respectively). **c**, All renditions from day 70 (a subset of the points in Fig. 2a). Colours correspond to repertoire time (50th percentile in Fig. 2d). Anticipations (repertoire times greater than 70) and regressions (repertoire times less than 70) occur at locations corresponding to vocalizations typical of later and earlier development (compare with Fig. 2a). Numbers 1–3 mark the approximate locations of the example renditions in **b**. **d**, Mixing matrices for additional birds (analogous to Fig. 2e, using the same birds as in Extended Data Fig. 2d). Bird 3 produced only a very few vocalizations (mostly calls) before tutoring onset (black arrows). The mixing matrices consistently show a period of gradual change starting after tutor onset and lasting several weeks. This gradual change typically slows down (resulting in larger mixing values far from the diagonal) at the end of the developmental period considered here (day 90 post-hatch; later periods are in Extended Data Fig. 6). Grey values correspond to the base-2 logarithm of the mixing ratio (LMR), that is, histograms over the pooled neighbourhood times (Fig. 2c) normalized by a null hypothesis obtained from a random distribution of production times (see Supplementary Methods). For example, an LMR value of 5 implies that renditions from the corresponding pair of production times are  $2^5 = 32$  times more mixed at the level of local neighbourhoods than would be expected by chance (that is, there is a random distribution of production times across renditions). **e**, As in **d**, bird 2, but after shuffling production times among all data points. Effects under this null hypothesis are small (the maximal observed mixing ratio is  $2^{0.06}$  or approximately 1.042). Similar, small effects under the null hypothesis are obtained for the other mixing matrices discussed throughout the text. **f–h**, Properties of the behavioural trajectory inferred from the mixing matrix in Fig. 2f. **f**, Pairwise distances between points along the inferred behavioural trajectory (x axis), plotted against measured disparities

(y axis). Disparities are obtained by rescaling and inverting the similarities in Fig. 2f (see Supplementary Methods). The points on the trajectory are inferred with ten-dimensional non-metric MDS on the measured disparities. Importantly, the pairwise distances between inferred points faithfully represent the corresponding, measured disparities (all points lie close to the diagonal; MDS stress = 0.0002). **g**, **h**, Structure of low-dimensional projections of the behavioural trajectory. We applied principle-component analysis to the ten-dimensional arrangement of points inferred with MDS and retained an increasing number of dimensions (number of dimensions indicated by greyscale). For example, the projection onto the first two principle components is shown in Fig. 2h (MDS dimension 2 in **g**, **h**). The first two principle components explain 75% of the variance in the full ten-dimensional trajectory. **g**, Measured (true) disparity (thick grey curve) and distances along the inferred trajectories (points and thin curves) as a function of the day gap ( $\delta$ ) between points. For any choice of projection dimensionality and  $\delta$ , we computed the Euclidean distances between any two points separated by  $\delta$  and averaged across pairs of points. The measured (true) disparities increase rapidly between subsequent and nearby days, but only slowly between far apart days (thick grey curve). Low-dimensional projections of the trajectory (for example, MDS dimension 2) underestimate the initial increase in disparities. **h**, Angle between the reconstructed direction of across-day change for inferred behavioural trajectories, as a function of the day gap between points. Same conventions and legend as in **g**. For the one- and two-dimensional trajectories, the direction of across-day change varies little or not at all from day to day (see inset; the arrow indicates the angle of across-day change). On the other hand, the direction of across-day change along the full, ten-dimensional behavioural trajectory is almost orthogonal for subsequent days. Data shown in **g**, **h** suggest that the full behavioural trajectory is more 'rugged' than indicated by the two-dimensional projection in Fig. 2h. This structure is consistent with the finding that across-day change includes a large component that is orthogonal to the directions of slow change and of within-day change (Fig. 3j). Note that **a** shows 200-ms spectrogram segments, whereas **b–h** are based on 68-ms segments (as are most of the analyses).

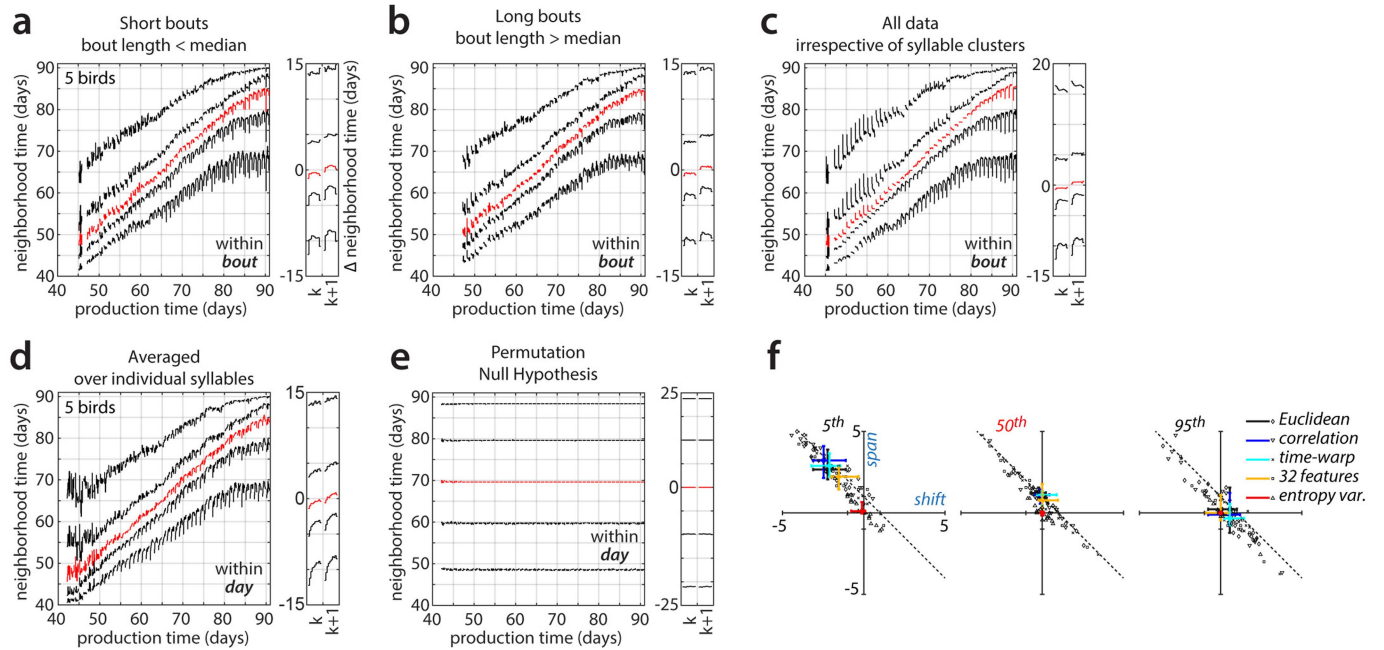


**Extended Data Fig. 4** | See next page for caption.



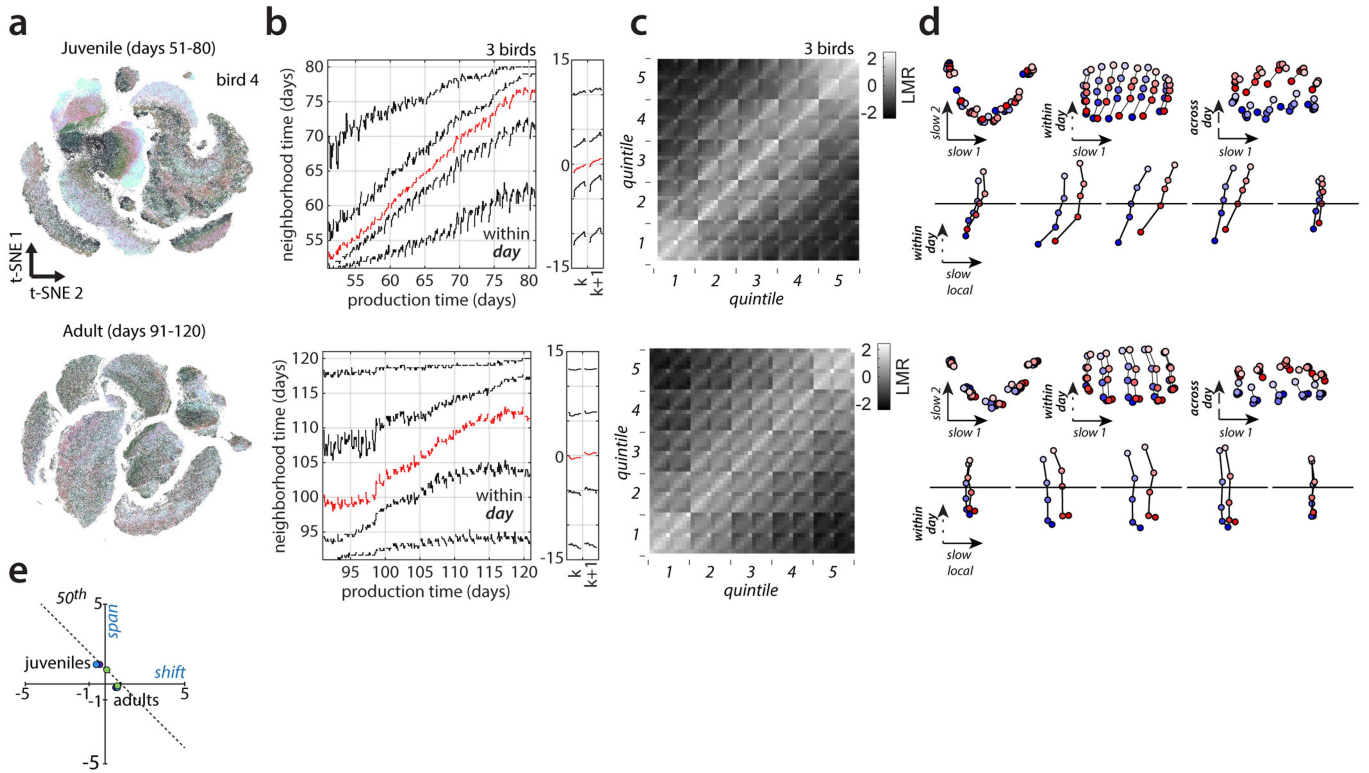
**Extended Data Fig. 4 | Models of the alignment between components of change. a–d,** Validation of repertoire dating. We simulated individual behavioural renditions as points in a high-dimensional space, drawn from a time-dependent probability distribution that changes both within and across days (see Supplementary Methods), and verified that repertoire dating can successfully recover the underlying structure of the models. The main parameters determining the relative alignment of the DiSC with the directions of within-day and across-day change are:  $k_a$ , the amount of within-day change along the DiSC;  $g_n$ , the amount of within-day change orthogonal to the DiSC; and  $b_n$ , the amount of across-day change orthogonal to the DiSC. These parameters are expressed relative to the amount of across-day change along the DiSC (thick black arrow in **a**, **b**). The two models shown (**a**, **c**) are characterized by different amounts of overnight consolidation of within-day changes along the DiSC. **a**, Model 1. Within-day change is aligned with the DiSC ( $g_n = 0$ ) and is large ( $k_a = 5$ ). The component of across-day change orthogonal to the DiSC is as large as the component of across-day change along it. In this scenario, overnight consolidation of within-day changes along the DiSC is weak for typical renditions (20% of change is consolidated, corresponding to a consolidation index of  $-0.8$  in Fig. 1c, d). **b**, Repertoire dating percentiles for model 1, analogous to Fig. 3b. The time course of the 50th repertoire-dating percentile (typical renditions, red) closely reproduces the dynamics of change along the DiSC implied by **a**: within-day change along the DiSC is large (the red line extends over about five days) and consolidation is weak (the starting point on day  $k + 1$  relative to day  $k$  moves by about 20% of the overall within-day range). **c**, Model 2. Within-day change has a large component orthogonal to the DiSC, whereas across-day change is aligned with the DiSC. In this scenario, overnight consolidation of within-day changes along the DiSC is strong (80% of the change is consolidated; consolidation index  $-0.2$ ) for typical renditions. **d**, Repertoire dating percentiles for model 2, analogous to **b**. The time course of the 50th repertoire dating percentile (typical renditions, red) closely reproduces the dynamics of change along the DiSC implied by **c**. In **b**, **d**,

differences between anticipations (95th percentile) and regressions (5th percentile) correctly reflect the underlying model parameters (see Supplementary Methods). **e–j**, Validation of stratified behavioural trajectories. We generated three sets of stratified behavioural trajectories that differ with respect to the alignment of within-day and across-day change with the DiSC. We built each set of trajectories by arranging 50 points (five strata per day, five production time periods per day, on two consecutive days; same conventions as Fig. 3f, g) within a four-dimensional space. We then generated simulated stratified mixing matrices (**e–g**, replotted from Fig. 3f) by computing pairwise distances between all points, and transforming distances into similarities. We visualize the behavioural trajectories (**h–j**) with the same two-dimensional projections as in Fig. 3h–j, with the same scale along all dimensions. In all models, overnight consolidation along the DiSC is perfect (strong consolidation) for all strata. **e**, Model 1: within-day change and across-day change occur only along the DiSC. For each stratum (that is, each of the five 10-by-10 squares along the diagonal), similarity decreases smoothly with time, reflecting the gradual progression of the trajectory along the DiSC within and across days. **f**, Model 2: within-day change has a large component that is not aligned with the DiSC. **g**, Model 3: both within-day and across-day change have large components that are not aligned with the DiSC. The misaligned component of across-day change reduces the similarity between day  $k$  and day  $k + 1$  compared with model 2, resulting in smaller values in the 5-by-5 squares comparing points from day  $k$  and day  $k + 1$ . **h**, Behavioural trajectories for model 1: the two-dimensional projection containing the DiSC (top) explains all the variance in the trajectories. **i**, Behavioural trajectories for model 2: similar to **h**, but points from different periods during the day are also displaced along an orthogonal direction of within-day change (middle). **j**, Behavioural trajectories for model 3: similar to **i**, but points from adjacent days are also displaced along an orthogonal direction of across-day change (bottom). Note that the models in **e–j** are implemented differently to the models in **a–d** (see Supplementary Methods).



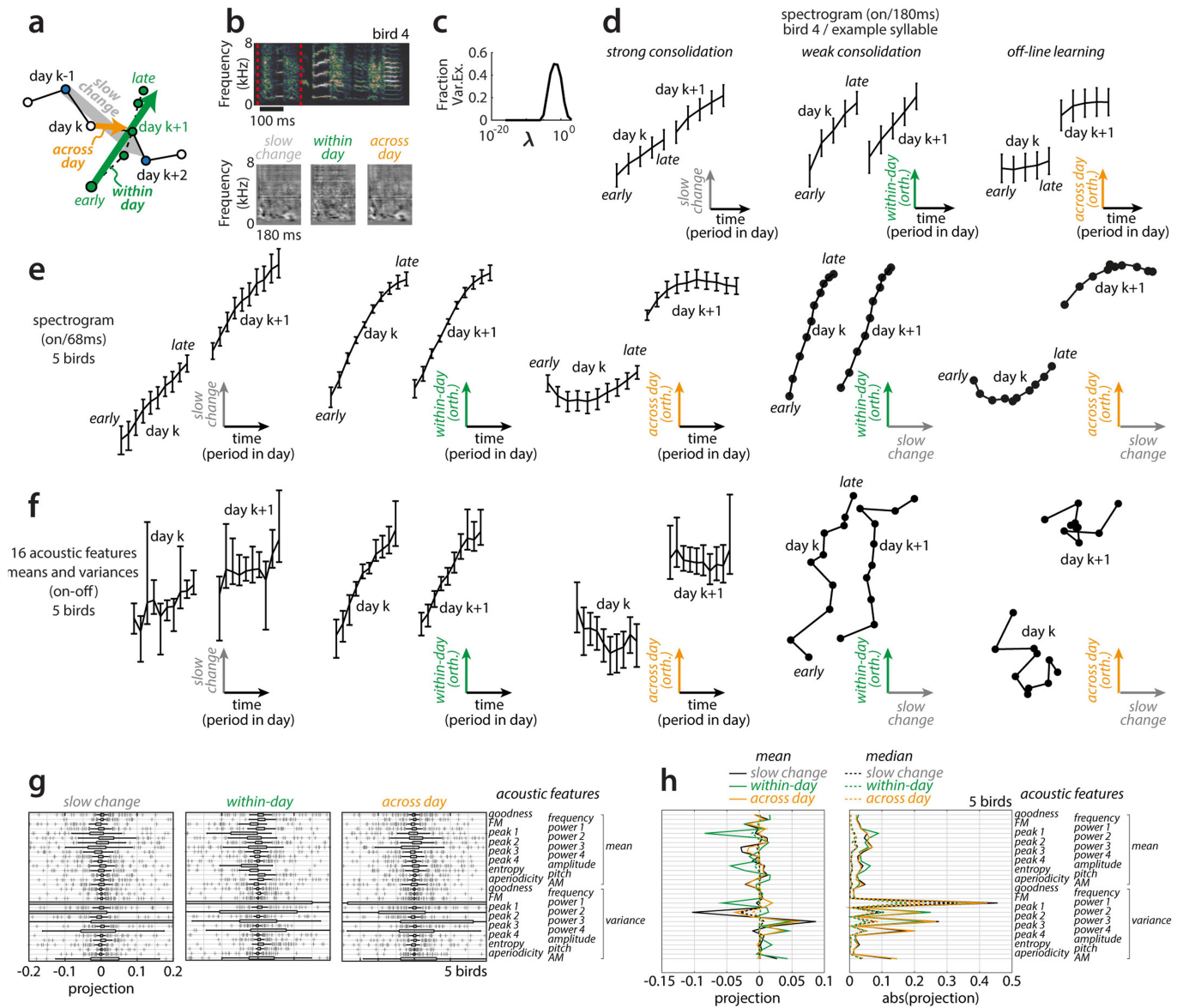
**Extended Data Fig. 5 | Repertoire dating control analyses. a–c.** Within-bout effects, analogous to Fig. 3c, d. **a**, Within-bout effects computed only from renditions that fall into short bouts (the bout length is less than the median). **b**, Analogous to **a**, but computed only from renditions that fall into long bouts (the bout length is more than the median). The changes in the behavioural repertoire observed within a bout are qualitatively similar for short and long bouts (compare **a**, **b**; within-bout effects are most pronounced after day 70). In particular, the song becomes more regressive shortly before the end of a bout (5th percentile, bottom curves). This suggests that the analogous effect in Fig. 3c, d occurs at the end of a bout, rather than at a fixed time after the beginning of a bout. **c**, Analogous to Fig. 3c, d but computed over the entire dataset without prior clustering into syllables. The changes in behavioural repertoire differ in several respects from those in Fig. 3c, d, which were computed on individual syllables and then averaged across syllables (see Supplementary Methods). Here, the increase in regressions at the bout end is less pronounced. Moreover, large within-bout changes also occur for anticipations early in development. Both differences may reflect changes in the relative frequency of renditions from each syllable (for example, introductory notes) sung throughout a bout. Such changes in frequency can affect the results in **c**, which were computed on the unclustered data, but not those in **a**, **b**, **d**. Within-day effects, analogous to Fig. 3a, b, but computed for individual syllables, and then averaged across syllables and animals. The

changes in behavioural repertoire are qualitatively similar to those in Fig. 3a, b, which were computed using the unclustered data. This similarity implies that the dynamics along the direction of slow change in Fig. 3 cannot be explained by changes during the day in the relative frequency of renditions from each syllable. **e**, Analogous to Fig. 3a, b but computed after shuffling production times among all data points. Within-day changes of the percentile curves are small under this null hypothesis. The maximal span of within-day fluctuations is 0.2 days, compared with 3.71 for the unshuffled data in Fig. 3b. The total repertoire spread (5th to 95th percentiles) is around 40 days, compared with around 23 days for unshuffled data. The 50th percentile curve is flat, implying that the shuffled data do not undergo a systematic drift over time (that is, do not describe a DiSC). The vertical separation between percentiles, then, reflects the range of production times in the data, not the spread along the DiSC. The time course of the 5th and 95th repertoire dating percentiles should thus be interpreted as the progression of regressions and anticipations along the DiSC only over the range of repertoire times covered by typical renditions (that is, approximately the vertical range of the 50th repertoire dating percentile). **f**, Analogous to Fig. 3e but for different distance metrics (Euclidean; correlation; Euclidean after time warping) and feature representations (32 acoustic features; 1 acoustic feature (entropy variance)). See also Extended Data Fig. 9.



**Extended Data Fig. 6 | Behavioural change in adult versus juvenile birds.**  
**a–d**, Comparison of within-day repertoire dating results during and after the end of development (averaged over three birds). Top, juvenile birds; bottom, same birds but as adults. **a**, Large-scale embeddings analogous to Fig. 2a. **b**, Repertoire dating percentiles, analogous to Fig. 3a, b. **c**, Stratified mixing matrix, analogous to Fig. 3g. **d**, Stratified behavioural trajectories, analogous to Fig. 3h–k. **e**, Shift and span values for the 50th percentile, for juvenile and adult birds. Points indicate individual birds. Song in adult birds is not static, but the time course of change differs from that observed in juveniles. First, change in adults is substantially less than in juveniles (see the slope of the 50th percentile in the top versus bottom parts of **b**). Second, the relation of fast (within-day) and slow (across-day) change differs in juveniles versus adults. In

juveniles, vocalizations move along the DiSC (y axes in **b**; slow local axis in **d**) within each day and the repertoire time of typical renditions increases by about one day from morning to evening (50th percentile; span is approximately one day) and is maintained through the next morning (shift is approximately 0 days). In adults, typical renditions do not show within-day progress along the DiSC (the span is approximately 0 days) but change overnight across days (the shift is greater than 0 days). In adults, the regressive tail of the repertoire in particular moves towards smaller values during the day (**b**, bottom right; 5th percentile), whereas in juveniles it consistently moves towards larger values (**b**, top). In both juvenile and adult birds, within-day change has a strong component that is misaligned with the DiSC (within-day axis in **d**).

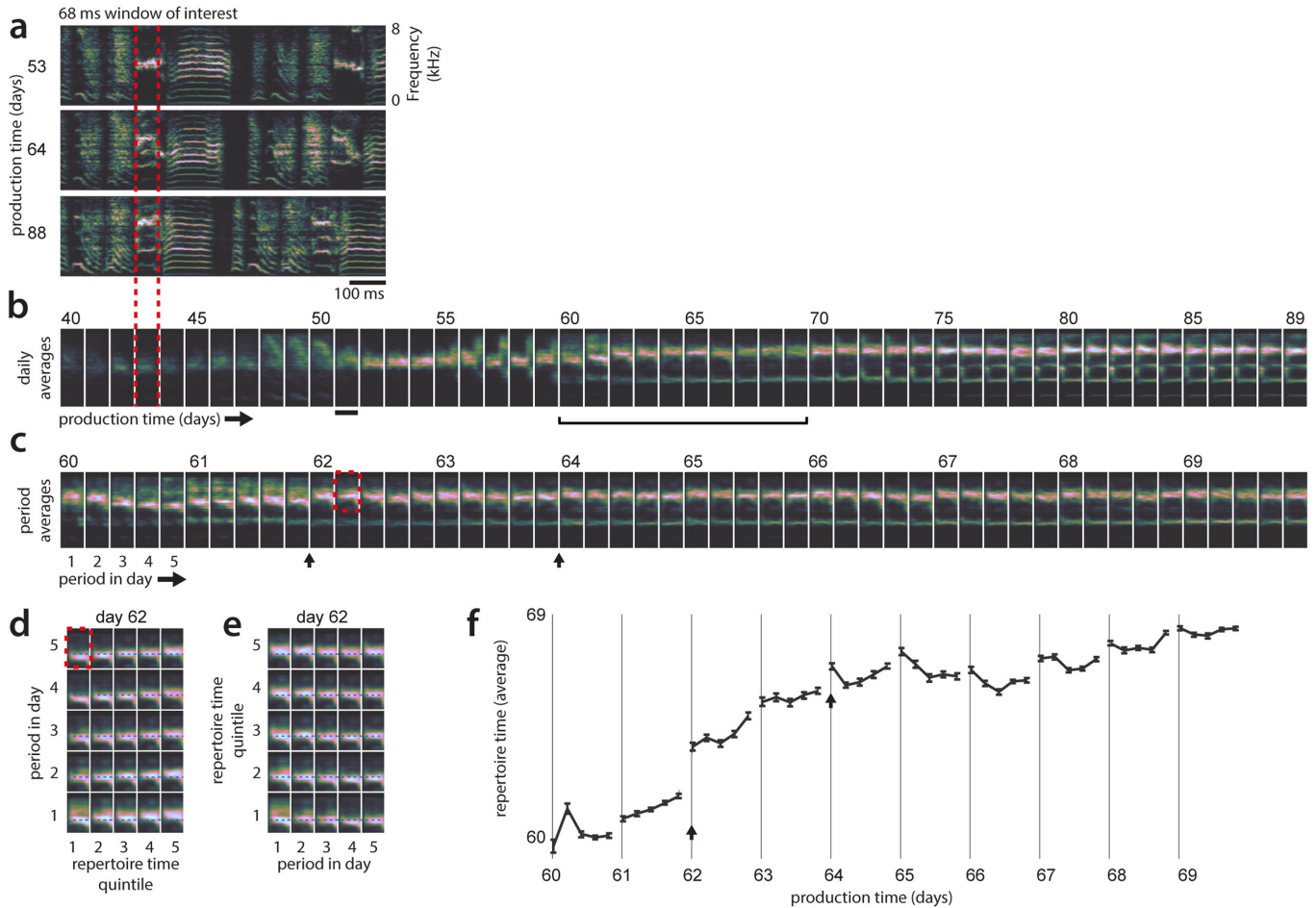


Extended Data Fig. 7 | See next page for caption.

**Extended Data Fig. 7 | Local linear analysis. a–e.** We validated the structure of change inferred with nearest-neighbour statistics (Fig. 3) with an approach based on linear regression in the high-dimensional spectrogram space (see Supplementary Methods). Unlike for the case of nearest neighbour-based statics, here each rendition must first be assigned to a cluster (that is, a syllable; compare with Fig. 2a) and each cluster is analysed separately. **a**, Illustration of the linearization scheme. First, we infer the (local) DiSC on days  $k$  and  $k + 1$  (grey arrow) as the vector of linear-regression coefficients relating production day to variability of renditions from days  $k - 1$  and  $k + 2$ . Second, we infer the direction of within-day change (green arrow) as the linear-regression coefficients relating the period within a day to variability of renditions from days  $k$  and  $k + 1$ , orthogonalized to the DiSC. Third, we infer the direction of across-day change (orange arrow) as the linear-regression coefficients relating production day to variability of renditions from days  $k$  and  $k + 1$ , orthogonalized to the DiSC and within-day change. All three sets of coefficients, and the corresponding directions in spectrogram space, typically vary across days, syllables and birds. The progression of song along the DiSC and along the (orthogonalized) directions of within-day and across-day change are obtained by projecting renditions on day  $k$  and  $k + 1$  onto the corresponding directions. **b**, Example rendition of syllable b as in Fig. 1 (top, encapsulated by red lines) and inferred coefficients (directions in spectrogram space; bottom) for day  $k = 57$ . Bright and dark shades of grey mark spectrogram bins for which power increases or decreases, respectively, over the corresponding timescales in **a**. **c**, Dependency of cross-validated regression quality (fraction of variance explained; y axis) on the regularization constant ( $\lambda$ ) for the estimation of the DiSC. One regularization constant was chosen for each syllable and the direction based on maximizing the leave-one-out cross-validation error on the training set. **d**, Progression of syllable b along the directions of change shown in **b**, during days 57 and 58. Renditions from each day are binned into ten consecutive periods on the basis of production time within the day (analogous to the ten periods in Fig. 3a, b; curves and error bars represent means and 95% bootstrapped confidence intervals). For simplicity of visualization, the time elapsed (x axis) during the night between days  $k$  and  $k + 1$  is not shown to scale. The position along the DiSC for the morning of day  $k + 1$  is close to that for the

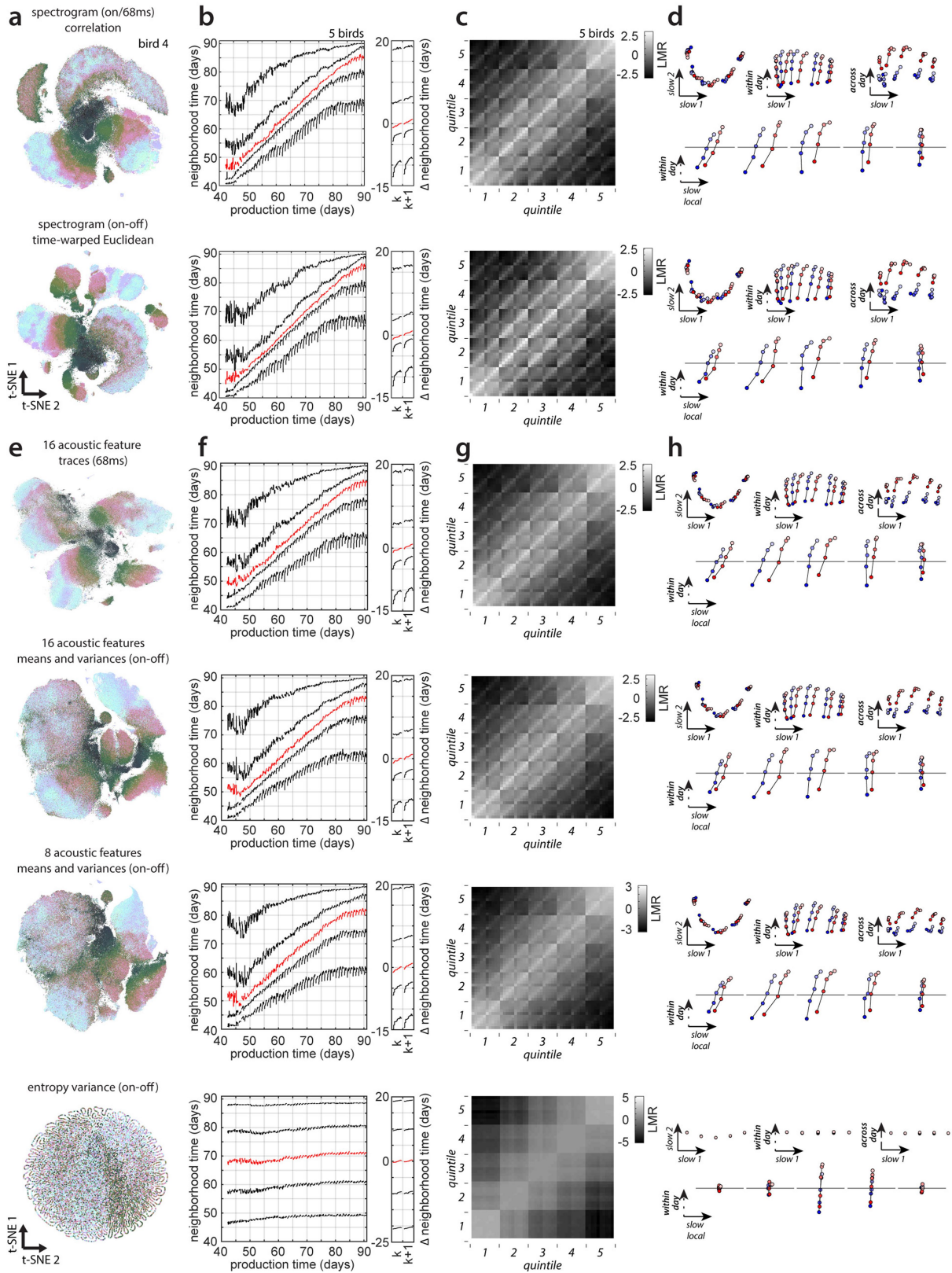
evening of day  $k$ , indicating overall strong consolidation (left). The position along the direction of within-day change is reset overnight, implying that the underlying changes are not consolidated (middle). The position along the direction of across-day change jumps overnight, consistent with offline learning (right). We note that strong consolidation, weak consolidation and offline learning have all been reported previously, albeit in different behaviours and species<sup>2,4,14,15,23,24</sup>. The charts in **d** show that these different patterns of change can occur in the very same syllable along distinct spectral features (see also Fig. 1h and Extended Data Fig. 8). By considering features with different projections onto these directions, a wide range of consolidation patterns can be uncovered (see also Fig. 1h). **e**, As for **d**, but averaged across all four-day windows during days 60–69 and over all syllables and birds (same five birds as in Figs. 2, 3). The resulting averages include contributions from the entire behavioural repertoire, including regressions, typical renditions and anticipations. The two right-most panels show concurrent progression along the DiSC and the direction of within-day or across-day change, combining data from the first and second, or first and third, panels in **e**. These representations are analogous, and in qualitative agreement, with the behavioural trajectories in Fig. 3h–k (typical). **f**, Analogous to **e**, but computed on vocalizations represented by 32 acoustic features instead of spectrograms. Directions as in **e** can be retrieved, but progression along the DiSC appears noisier, suggesting that the 32 acoustic features do not fully capture in particular the slow spectral changes occurring over development (see also Extended Data Fig. 9). **g**, **h**, Contribution of individual acoustic features to the directions of slow, within-day and across-day change. As in **f**, the directions are computed in the space of 32 acoustic features. **g**, Distribution of coefficients in the retrieved orthonormalized directions. Thick and thin black bars represent means and 95% confidence intervals; crosses show outliers; thin vertical lines represent medians. **h**, Means (solid lines) and medians (dotted lines) of the signed (left) or unsigned (right) distributions in **g**. Most coefficients are small and variable, indicating that the alignment between any of the 32 acoustic features and the inferred directions of change is weak and highly variable over time, syllables and birds.





**Extended Data Fig. 8 | Behavioural variability and stratification in an example syllable.** **a**, Songs of an example bird for three days during development. Only those spectrogram segments that belong to a particular syllable and location in the motif (68-ms window of interest; red dotted lines) are analysed in the subsequent panels. **b**, Developmental changes over the course of weeks. Renditions are binned by production day, and averaged. The most apparent changes are an increase in pitch and the later successive appearance of additional spectral lines at low frequencies. **c**, Within-day and across-day changes for days 60–69. Renditions are binned into five production-time periods spanning a day and averaged within bins. On many days, the changes within a day do not appear to recapitulate the changes occurring across days (for example, days 60 and 65; within-day progression does not smoothly transition between the vocalizations on preceding and subsequent days; see **b**). The averages also reveal occasional overnight ‘jumps’ in the properties of the vocalizations (see, for example, the vertical black arrows).

**d**, Comparison of within-day change and change on longer timescales. Renditions within each period and day were split into strata according to their repertoire times (for example, the quintiles in Fig. 2d), resulting in 25 averages, one for each combination of stratum and period within the day. Only the upper part of the spectrogram is shown (red rectangle in **c**). The progression along strata (*x* axis) emphasizes the large extent of motor variability along the DiSC existing within a single day (day 62). **e**, Same averages as in **d**, but with *x* and *y* axes swapped. In particular for regressive renditions (quintile 1), change within day 62 (*x* axis) does not recapitulate developmental changes occurring over months (*x* axis in **d**). **f**, Repertoire dating based on repertoire time (as in Extended Data Fig. 3c). Each point corresponds to a production-time period and the average of all repertoire times of renditions in that period. Error bars show bootstrapped 95% confidence intervals. The change in repertoire time, which is computed without using a low-dimensional parametrization of vocalizations captures the movement along the DiSC.



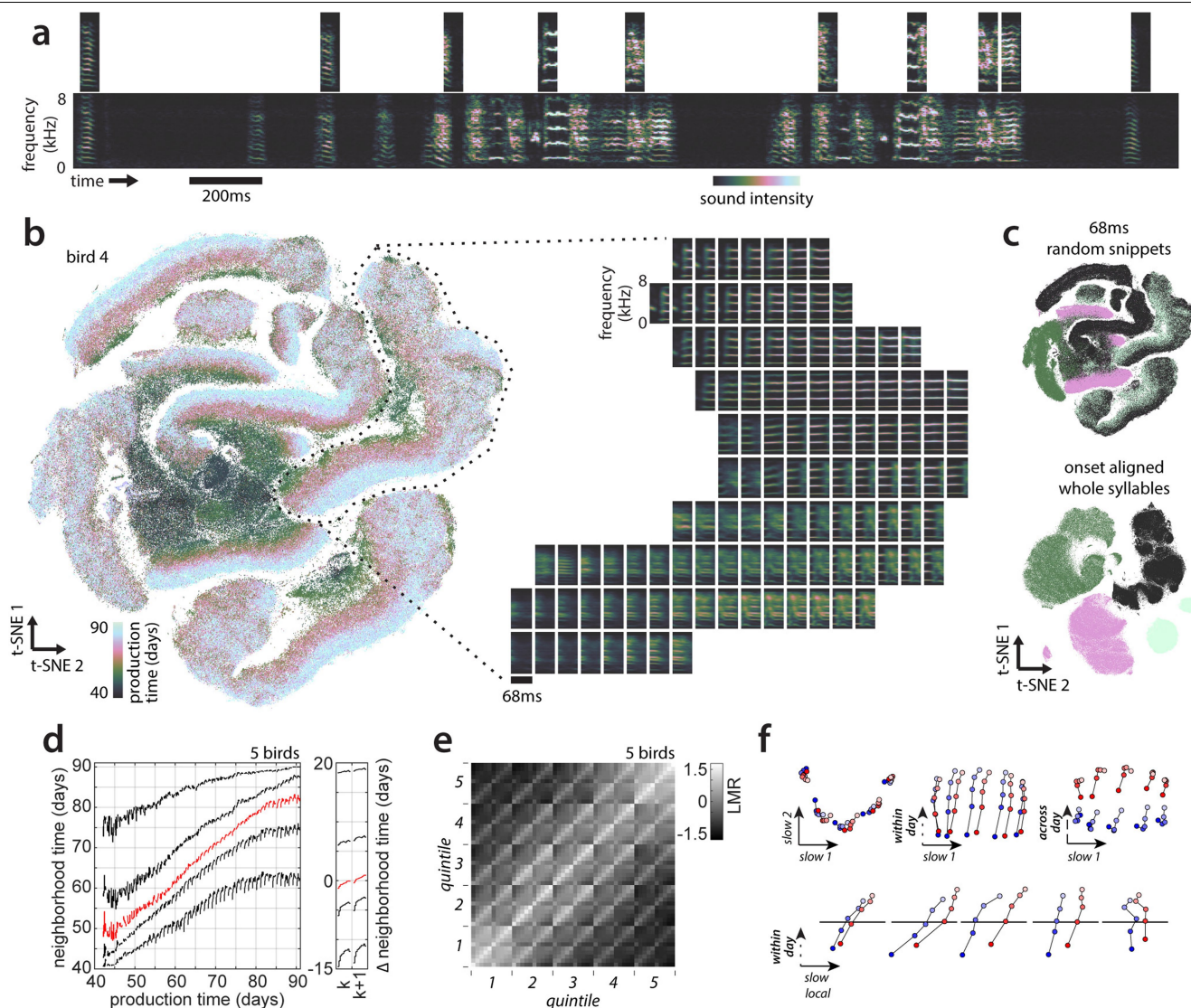
**Extended Data Fig. 9** | See next page for caption.

# Article

**Extended Data Fig. 9 | Behavioural change based on alternative distance metrics and features.** To demonstrate the robustness of the proposed nearest-neighbour statistics, we verified that the inferred time course of behavioural change is reproducible using a number of different distance metrics (used to define nearest neighbours) and parameterizations of vocalizations. **a–d**, We recomputed the main analyses using a Pearson's correlation metric on 68-ms onset-aligned spectrogram segments (first row); and the Euclidean distance on onset-to-offset spectrogram segments that were linearly time-warped to a duration of 100 ms (second row). For comparison, the main analyses in the text were based on Euclidean distance on 68-ms onset-aligned spectrogram segments (for example, Fig. 2c–f, 3). **a**, *t*-SNE visualization based on the corresponding distance metrics and sound representation for the example bird, analogous to Fig. 2a. **b**, Repertoire dating averaged over birds, analogous to Fig. 3a, b. **c**, Stratified mixing matrices averaged over birds, analogous to Fig. 3g. The mixing values are highly correlated across distance metrics: Euclidean (main text) versus correlation, variance explained = 92%; Euclidean (main text) versus time-warped Euclidean, 93%. **d**, Stratified behavioural trajectories based on **c**, analogous to Fig. 3h–k. The results in **a–d** are consistent with those in Fig. 3, showing that our findings are robust with respect to the exact definition of nearest neighbours. Moreover, the overall structure of the behavioural trajectory appears to depend only minimally on changes in tempo and spectrogram magnitude (first row: Pearson's correlation is invariant to changes in overall magnitude of vocalizations; second row: time-warped Euclidean distance is invariant to changes in tempo). **e–h**, We recomputed all main analyses with four additional parameterizations of vocalizations: time-dependent normalized acoustic feature traces for 16 acoustic features within 68-ms windows after syllable onset (first row); means and variances of the same 16 acoustic features over entire syllables (second row); means and variances of 8 of the 16 acoustic features (third row); and a one-dimensional parametrization consisting solely of entropy variance computed over entire syllables (fourth row). Feature means and variances were z-scored across all syllables. For all of these parameterizations we defined nearest

neighbours with the Euclidean distance. **e**, Embedding using *t*-SNE based on the corresponding parameterization and metric. For entropy variance alone, the embedding appears locally one dimensional (for visibility, data points are larger than for the other parameterizations). Entropy variance maps mostly smoothly onto this one-dimensional manifold (data not shown). **f**, Repertoire dating averaged over birds, analogous to Fig. 3a, b. Repertoire dating based on entropy variance alone fails to reproduce most of the results in Fig. 3 obtained with spectrogram segments. The percentile curves are almost flat, indicating that renditions cannot be reliably assigned to their production times on the basis of entropy variance alone. In this case, vertical separation between percentiles cannot be interpreted as spread along the DiSC (see Extended Data Fig. 5e). For entropy variance alone, span is greater than zero across all percentiles, but consolidation is consistently close to zero. **g**, Stratified mixing matrix averaged over birds, analogous to Fig. 3g. The match with the mixing matrix in Fig. 3g decreases as the dimensionality of the parameterization is reduced (spectrogram versus time-dependent feature traces: variance explained = 93%; spectrogram versus 16 acoustic feature means and variances, 91%; spectrogram versus 8 acoustic feature means and variances, 84%; spectrogram versus entropy variance, 54%). **h**, Stratified behavioural trajectories based on **g**, as in Fig. 3h–k. The inferred behavioural trajectories are similar across the first three song parameterizations. However, these alternative parameterizations result in more vertical separation between percentiles in **f**, suggesting that they capture the direction of slow change less well (compare with Fig. 3a and Extended Data Fig. 5e). Parameterizations of reduced dimensionality also result in progressively less defined syllable clusters in the embeddings (**e**, top to bottom). These observations suggest that a parameterization based on the full spectrogram is better suited to capture the different directions of change explored during development (see also Extended Data Fig. 7). Note that for entropy variance (bottom row), the projections onto the local direction of slow change are highly magnified compared with the projections in the top panels.





**Extended Data Fig. 10 | Behavioural change based on random spectrogram segments.** We recomputed all main-text analyses with a random segmentation of behaviour that does not require alignment to syllable onsets. This segmentation scheme can be applied to behaviour that does not fall into temporally discrete elements. Here each data point corresponds to a randomly chosen 68-ms spectrogram snippet drawn from a period of singing. Not all song was sampled, as we used 1,000,000 non-overlapping segments for each bird (see Supplementary Methods). **a**, Vocalizations of the example bird (as in Fig. 1a) from day 76, with example segments used for the analysis (at the top). **b**, t-SNE visualization for random segments from the example bird, based on nearest neighbours defined with respect to the Euclidean distance (left) and average spectrograms for different locations in the (t-SNE) embedding (right; analogous to Fig. 2b). Clusters corresponding to individual syllables are elongated compared with Fig. 2a. Variation along one direction within the cluster tends to account for production time (colour bar), while variation along another direction tends to reflect the timing of segments relative to syllable onsets. **c**, Embedding from Fig. 2a (bottom) and embedding of random 68-ms segments (top). Points in both embeddings are coloured according to cluster

identity defined on onset-aligned spectrogram segments covering entire syllables (bottom). The colour of each point corresponding to a random snippet (top) corresponds to the cluster identity of the surrounding syllable. Some clusters in the embedding based on random segments contain points assigned to two different syllables (for example, black versus green colours). **d**, Repertoire dating averaged over birds, analogous to Fig. 3a, b, e. Stratified mixing matrix averaged over birds, analogous to Fig. 3g. The mixing values are highly correlated with those in Fig. 3g (variance explained = 89%). **f**, Stratified behavioural trajectories based on e, as in Fig. 3h-k. The results in **d-f** largely reproduce the corresponding findings obtained with onset-aligned 68-ms spectrogram segments (Fig. 3) as well as with other song parameterizations (Extended Data Fig. 9). Nonetheless, the overall effect sizes are reduced, probably because of the additional variability introduced by the random position of segments relative to syllable onsets. In **d**, the vertical separation between the 5th and 95th percentiles is increased and the slope of 50th percentile is reduced compared with the main-text analyses (Fig. 3a), suggesting a noisier representation of the direction of slow change (see Extended Data Fig. 5e) compared with onset-aligned 68-ms segments (Fig. 3).

## Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- ☐ ☒ The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- ☐ ☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- ☐ ☒ The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- ☐ ☒ A description of all covariates tested
- ☐ ☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- ☐ ☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- ☒ ☐ For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give  $P$  values as exact values whenever suitable.*
- ☒ ☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- ☒ ☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- ☐ ☒ Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

Data collection Labview, Matlab

Data analysis Matlab

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

### Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

### Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- ☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)



## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	For each animal all available data in the given developmental window of interest was chosen.
Data exclusions	Out of a set of 16 animals, 11 were excluded because recording conditions were inadvertently changed mid-experiment by modifying the amplifier settings.
Replication	We did not attempt replication using a different dataset.
Randomization	n/a
Blinding	Describe whether the investigators were blinded to group allocation during data collection and/or analysis. If blinding was not possible, describe why OR explain why blinding was not relevant to your study.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems		Methods	
n/a	Involved in the study	n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies	<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines	<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology	<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging
<input type="checkbox"/>	<input checked="" type="checkbox"/> Animals and other organisms		
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants		
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data		

## Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals	Zebra Finches, Male
Wild animals	Provide details on animals observed in or captured in the field; report species, sex and age where possible. Describe how animals were caught and transported and what happened to captive animals after the study (if killed, explain why and describe method; if released, say where and when) OR state that the study did not involve wild animals.
Field-collected samples	For laboratory work with field-collected samples, describe all relevant parameters such as housing, maintenance, temperature, photoperiod and end-of-experiment protocol OR state that the study did not involve samples collected from the field.
Ethics oversight	All experimental procedures were approved by the Veterinary Office of the Canton of Zurich.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

# Decoding the development of the human hippocampus

<https://doi.org/10.1038/s41586-019-1917-5>

Received: 4 April 2019

Accepted: 12 November 2019

Published online: 15 January 2020

Suijuan Zhong<sup>1,10</sup>, Wenyu Ding<sup>2,10</sup>, Le Sun<sup>1,3,4,10</sup>, Yufeng Lu<sup>1,4,10</sup>, Hao Dong<sup>1,4</sup>, Xiaoying Fan<sup>5</sup>, Zeyuan Liu<sup>1,4</sup>, Ruiguo Chen<sup>1,4</sup>, Shu Zhang<sup>5</sup>, Qiang Ma<sup>1,4</sup>, Fuchou Tang<sup>5,6,7</sup>, Qian Wu<sup>2,8\*</sup> & Xiaoqun Wang<sup>1,3,4,9\*</sup>

The hippocampus is an important part of the limbic system in the human brain that has essential roles in spatial navigation and the consolidation of information from short-term memory to long-term memory<sup>1,2</sup>. Here we use single-cell RNA sequencing and assay for transposase-accessible chromatin using sequencing (ATAC-seq) analysis to illustrate the cell types, cell lineage, molecular features and transcriptional regulation of the developing human hippocampus. Using the transcriptomes of 30,416 cells from the human hippocampus at gestational weeks 16–27, we identify 47 cell subtypes and their developmental trajectories. We also identify the migrating paths and cell lineages of PAX6<sup>+</sup> and HOPX<sup>+</sup> hippocampal progenitors, and regional markers of CA1, CA3 and dentate gyrus neurons. Multiomic data have uncovered transcriptional regulatory networks of the dentate gyrus marker PROX1. We also illustrate spatially specific gene expression in the developing human prefrontal cortex and hippocampus. The molecular features of the human hippocampus at gestational weeks 16–20 are similar to those of the mouse at postnatal days 0–5 and reveal gene expression differences between the two species. Transient expression of the primate-specific gene *NBPFI* leads to a marked increase in PROX1<sup>+</sup> cells in the mouse hippocampus. These data provides a blueprint for understanding human hippocampal development and a tool for investigating related diseases.

The hippocampal formation (hippocampus) is a compound structure under the cerebral cortex in primates that forms and stores long-term memory by consolidating information from short-term memory, and also processes spatial information and navigation<sup>1,2</sup>.

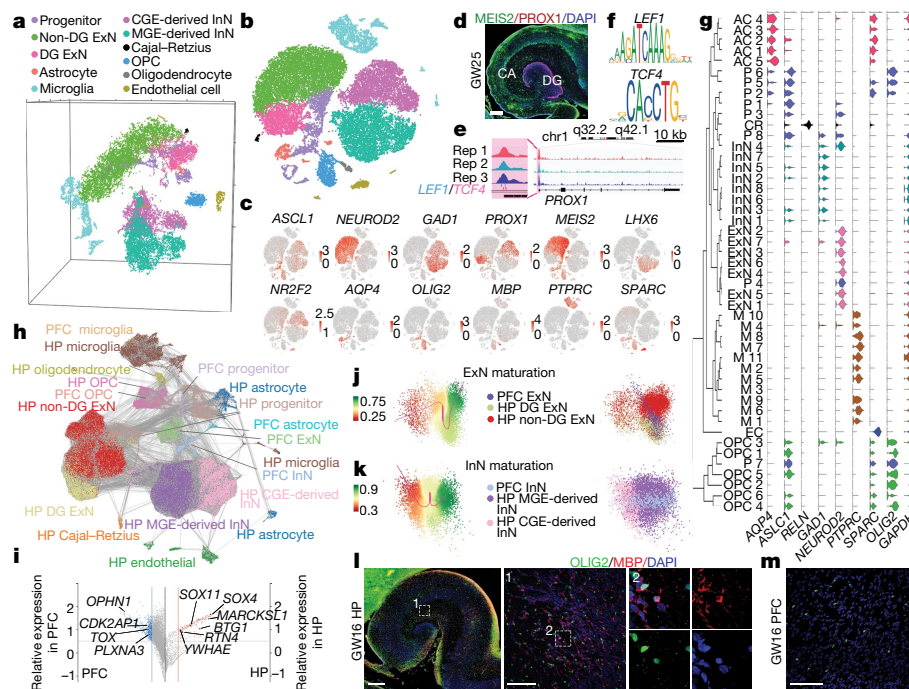
## Hippocampus single-cell transcriptome

To understand the molecular features of hippocampal cells during human brain development, we analysed 30,416 cells from the entire left hippocampus (including the hippocampus proper, the dentate gyrus (DG) and some of the subiculum connected to the hippocampus proper) at gestational weeks (GW) 16–27 (Supplementary Table 1) by droplet-based single-cell RNA sequencing (scRNA-seq). We performed *t*-distributed stochastic neighbour embedding (*t*-SNE) analysis and identified cells as progenitors, excitatory neurons (ExN), inhibitory neurons (InN), Cajal Retzius cells, astrocytes, oligodendrocyte progenitor cells (OPCs), oligodendrocytes, microglia and endothelial cells by using classic markers and gene ontology (GO) of differentially expressed genes (DEGs) (Fig. 1a–c, Extended Data Fig. 1a–c). The distributions of samples from two individuals at GW22 were similar on the

*t*-SNE plot (Extended Data Fig. 1d). We then used the DG marker *PROX1* to subclassify the ExN as DG ExN or non-DG ExN. The InN were further subclassified as being derived from the medial or caudal ganglionic eminence (MGE or CGE) on the basis of *LHX6* and *NR2F2* expression (Fig. 1a–d). *PROX1* is an essential transcription factor for the genesis of hippocampal granule cells and formation of the DG<sup>3,4</sup>. By searching transcription factor motifs identified from ATAC-seq peaks close to the *PROX1* transcription start site (TSS), we found three potential binding sites for LEF1 or TCF4, indicating that WNT signals are crucial for the production of DG granule cells (Fig. 1e, f), which is consistent with reported studies<sup>5,6</sup>. We further segregated cells into 47 distinct hierarchical subtypes by principal component analysis (PCA), showing that different subtypes of progenitors were highly correlated with fate-determined cells (Fig. 1g, Extended Data Fig. 2a–c).

To study developmental differences between the hippocampus and neocortex, we compared the transcriptome of the hippocampus (GW16–27) with that of the human prefrontal cortex (PFC) (GW8–26)<sup>7</sup> (Fig. 1h) and found differences in gene expression between the PFC and hippocampus across all cell types (Fig. 1i, Supplementary Table 2). The HMG box domain-containing protein TOX was highly expressed

<sup>1</sup>State Key Laboratory of Brain and Cognitive Science, CAS Center for Excellence in Brain Science and Intelligence Technology, Institute of Brain-Intelligence Technology (Shanghai), Institute of Biophysics, Chinese Academy of Sciences, Beijing, China. <sup>2</sup>State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing, China. <sup>3</sup>Institute for Stem Cell and Regeneration, Chinese Academy of Sciences, Beijing, China. <sup>4</sup>University of Chinese Academy of Sciences, Beijing, China. <sup>5</sup>Beijing Advanced Innovation Center for Genomics, College of Life Sciences, Peking University, Beijing, China. <sup>6</sup>Biomedical Institute for Pioneering Investigation via Convergence and Center for Reproductive Medicine, Ministry of Education Key Laboratory of Cell Proliferation and Differentiation, Beijing, China. <sup>7</sup>Peking-Tsinghua Center for Life Sciences, Peking University, Beijing, China. <sup>8</sup>IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing, China. <sup>9</sup>Beijing Institute for Brain Disorders, Beijing, China. <sup>10</sup>These authors contributed equally: Suijuan Zhong, Wenyu Ding, Le Sun, Yufeng Lu. \*e-mail: qianwu@bnu.edu.cn; xiaoqunwang@ibp.ac.cn



**Fig. 1 | Molecular diversity of single cells from the developing human hippocampus. a–c.** Visualization of eleven major classes using t-SNE in 3D (a) and 2D (b) visualization. c, Expression of known markers. HP, hippocampus; PFC, prefrontal cortex. Dots, individual cells; grey, no expression; red, relative expression (log-normalized gene expression). d, Immunostaining of MEIS2 and PROX1. Scale bar, 500  $\mu$ m. e, Normalized ATAC-seq profiles of *PROX1* in GW25 hippocampus show the activation of *PROX1*. Amplified view (pink) shows predicted LEF1 and TCF4 binding sites. f, LEF1 and TCF4 binding motifs are identified in the ATAC-seq peaks close to the *PROX1* TSS. g, Hierarchical clustering analysis of 47 subclasses. AC, astrocyte; P, progenitor; CR, Cajal–Retzius cell; M, microglia; EC, endothelial cell.  $n = 134, 141, 95, 275, 58, 300, 397, 159, 204, 483, 101, 74, 670, 1,019, 1,765, 2,334, 793, 1,073, 909, 3,189, 2,347, 92,$

1,838, 717, 1,956, 730, 1,192, 2,573, 54, 259, 84, 84, 44, 489, 465, 246, 68, 229, 638, 540, 131, 103, 139, 257, 227, 459 and 282 cells, top to bottom. h, Abstracted graph shows the connections on the transcriptome between different cell types in the developing human hippocampus and PFC. i, Scatterplot of all genes for correlation with conserved differentiation network across PFC and hippocampus. Blue plot shows genes related to PFC; red plot shows genes related to hippocampus. j, k, Maturation scores of excitatory neurons (j) and inhibitory neurons (k) in PFC and hippocampus. l, m, Immunostaining for oligodendrocyte markers at GW16 in human hippocampus and prefrontal cortex. Scale bars, 500  $\mu$ m (l, left); 100  $\mu$ m (l, right, m). The experiment was repeated three times independently with similar results.

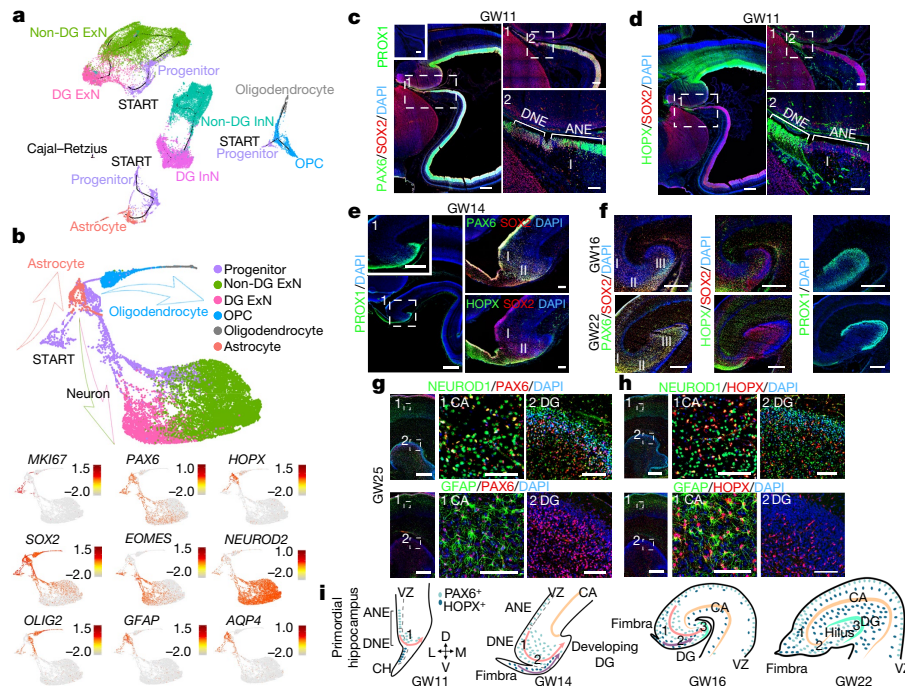
in the PFC, whose progenitors are regulated by HMGA2<sup>28</sup>. SOX4 and SOX11, two SOXC transcription factors that are required for neuronal differentiation during neurogenesis in the adult hippocampus<sup>9,10</sup>, were relatively highly expressed in the hippocampus (Fig. 1i). GO analysis of DEGs between ExN from the hippocampus and PFC at GW16 indicate that hippocampal ExN may undergo synapse organization and axonogenesis at GW16 (Extended Data Fig. 2d). Comparison of the maturation trajectories of hippocampus and PFC neurons indicated that hippocampus non-DG ExN were more mature than PFC ExN, whereas maturation of DG ExN was similar to that of PFC ExN (Fig. 1j). InN of the PFC and hippocampus generally showed a similar maturation status, whereas MGE-derived InN were more mature than CGE-derived InN in the hippocampus (Fig. 1k). Consistent with transcriptome analysis, immunofluorescence staining for OLIG2 and MBP showed a number of MBP<sup>+</sup> cells in the subfield of the hippocampus, whereas no MBP<sup>+</sup> cells were found in the human PFC at GW16 (Fig. 1l, m, Extended Data Fig. 2e), suggesting that oligodendrocytes may be involved the maturation of hippocampal neurons during early development.

## Progenitors of the developing hippocampus

To further investigate cellular lineage relationships in the fetal human hippocampus, we reconstructed five developmental paths by monocle analysis without microglia and endothelial cells (Fig. 2a). Three major subgroups of progenitors differentiated to excitatory neuronal, OPC and oligodendrocyte or astrocyte lineages. The MGE- and CGE-derived InN were separated in different directions, which is consistent with previous studies showing that hippocampal InN originate from different

progenitors located in the ganglionic eminence<sup>11</sup>. To further reveal the diversity and molecular properties of human hippocampal progenitors, we used GO analysis of DEGs and marker genes to identify eight subclusters (Extended Data Fig. 3a–c). *EOMES*<sup>+</sup>, *MEIS2*<sup>+</sup> and *NEUROD2*<sup>+</sup> progenitors were in clusters P3 and P4, indicative of ExN generation (Extended Data Fig. 3a). *AQP4*, *OLIG1/OLIG2* and *PDGFRA* were highly expressed in clusters P5, P6 and P7, respectively, indicative of astrocyte and oligodendrocyte cell fates. Cluster P8 contained a small number of progenitors that highly expressed *DLX1* and *DLX2*, indicating that these cells may differentiate as InN (Extended Data Fig. 3a–f).

To understand how progenitors develop into neuronal and glial cells, we carried out trajectory analysis (Fig. 2b) and separated three paths towards neurons, astrocytes and oligodendrocytes. Notably, PAX6<sup>+</sup> and HOPX<sup>+</sup> progenitors, which are considered as neurogenic progenitors in the neocortex<sup>12</sup>, were likely to contribute to both neurogenesis and gliogenesis in the human hippocampus (Fig. 2b, Extended Data Fig. 3g, h). We next examined the locations of cells expressing PAX6 or HOPX by immunofluorescence staining (Fig. 2c–f). At GW11, the primordial hippocampal area, located adjacent to the cortical hem (CH), was composed of the dentate neuroepithelium (DNE) and ammonic neuroepithelium (ANE). The majority of cells in the DNE and ANE expressed SOX2, and PAX6<sup>+</sup> SOX2<sup>+</sup> progenitors of ANE started to migrate (Fig. 2c, Extended Data Fig. 4a). At the same time, HOPX<sup>+</sup> SOX2<sup>+</sup> DNE progenitors also indicated migration potential (Fig. 2d, Extended Data Fig. 4b). As the hippocampus developed at GW14, a number of PAX6<sup>+</sup> progenitors migrated away from the ventricular zone towards the future DG (PROX1<sup>+</sup> region, Fig. 2e, Extended Data Fig. 4c), forming the primary matrix (I) and secondary matrix (II). HOPX<sup>+</sup> progenitors also migrated in the same



**Fig. 2 | Molecular signature of neural progenitor cells of the developing human hippocampus.** **a**, Cell lineage relationships of all cells analysed except for microglia and endothelial cells in developing human hippocampus. Monocle recovered a branched single-cell trajectory beginning with progenitors and terminating at excitatory neurons, inhibitory neurons, astrocytes and oligodendrocytes. **b**, Cell lineage relationships of progenitors, excitatory neurons, astrocytes and oligodendrocytes in developing human hippocampus. Known gene expression is shown below. Arrows show the directions of lineages. **c**, Immunofluorescence images of PROX1 (scale bar, 200  $\mu$ m), PAX6 and SOX2 at GW11. Scale bars, 500  $\mu$ m (left), 200  $\mu$ m (top right), 100  $\mu$ m (bottom right). **d**, Immunofluorescence images of

HOPX and SOX2 at GW11. Scale bars, 500  $\mu$ m (left), 200  $\mu$ m (top right), 100  $\mu$ m (bottom right). **e**, Immunofluorescence images of PROX1 (scale bars, 1,000  $\mu$ m, inset 500  $\mu$ m), PAX6, HOPX and SOX2 in GW14. Scale bar, 200  $\mu$ m. **f**, Immunofluorescence images of PROX1 (scale bar, 1,000  $\mu$ m), PAX6, HOPX and SOX2 at GW16 (top) and GW22 (bottom). Scale bar, 500  $\mu$ m. **g**, Immunofluorescence images of PAX6, HOPX, NEUROD1 and GFAP at GW25. Scale bars, 500  $\mu$ m (left); 100  $\mu$ m (right). **h**, The experiments were repeated three times independently with similar results. **i**, Schema depicting locations of PAX6<sup>+</sup> or HOPX<sup>+</sup> progenitors in developing human hippocampus from GW11 to GW22. Arrows indicate direction of migration.

direction but closer to the pial side (Fig. 2e, Extended Data Fig. 4c). At GW16, the migration of PAX6<sup>+</sup> and HOPX<sup>+</sup> progenitors continued and many cells arrived at the hilus and formed an origin hub of DG cells, called the tertiary matrix (III). Notably, PAX6<sup>+</sup> progenitors were located outside HOPX<sup>+</sup> progenitors while migrating (Fig. 2f, Extended Data Fig. 4d–f). PAX6<sup>+</sup> progenitors were still abundant, and some were located in the blades of the DG, but only some HOPX<sup>+</sup> progenitors were found in the hilus; the majority of HOPX<sup>+</sup> progenitors were in the cornu ammonis (CA) at GW 22 (Fig. 2f, Extended Data Fig. 4g–i).

We next evaluated the proliferation capacity and cell fate of PAX6<sup>+</sup> and HOPX<sup>+</sup> progenitors. Both scRNA-seq data and immunostaining indicate that a subpopulation of PAX6<sup>+</sup> and HOPX<sup>+</sup> progenitors are active in the cell cycle even at the mid-gestational stage (Extended Data Fig. 5a–d). In cell fate assessment, we observed PAX6<sup>+</sup> NEUROD1<sup>+</sup> cells in the CA and DG, but PAX6<sup>+</sup> GFAP<sup>+</sup> cells only in the CA (Fig. 2g, Extended Data Fig. 5e–g). Similar expression patterns were found in HOPX<sup>+</sup> cells (Fig. 2h, Extended Data Fig. 5h, i). Next, we evaluated the maturation status of PAX6<sup>+</sup> or HOPX<sup>+</sup> progenitors and found that NEUROD1<sup>+</sup> cells were more mature than GFAP<sup>+</sup> cells, suggesting that they may have been born earlier (Extended Data Fig. 5j, k). Together, our data suggest that although the origins and migrating paths of PAX6<sup>+</sup> and HOPX<sup>+</sup> progenitors differ, they both contribute to neural and glial genesis in a spatiotemporal manner in the developing human hippocampus (Fig. 2i).

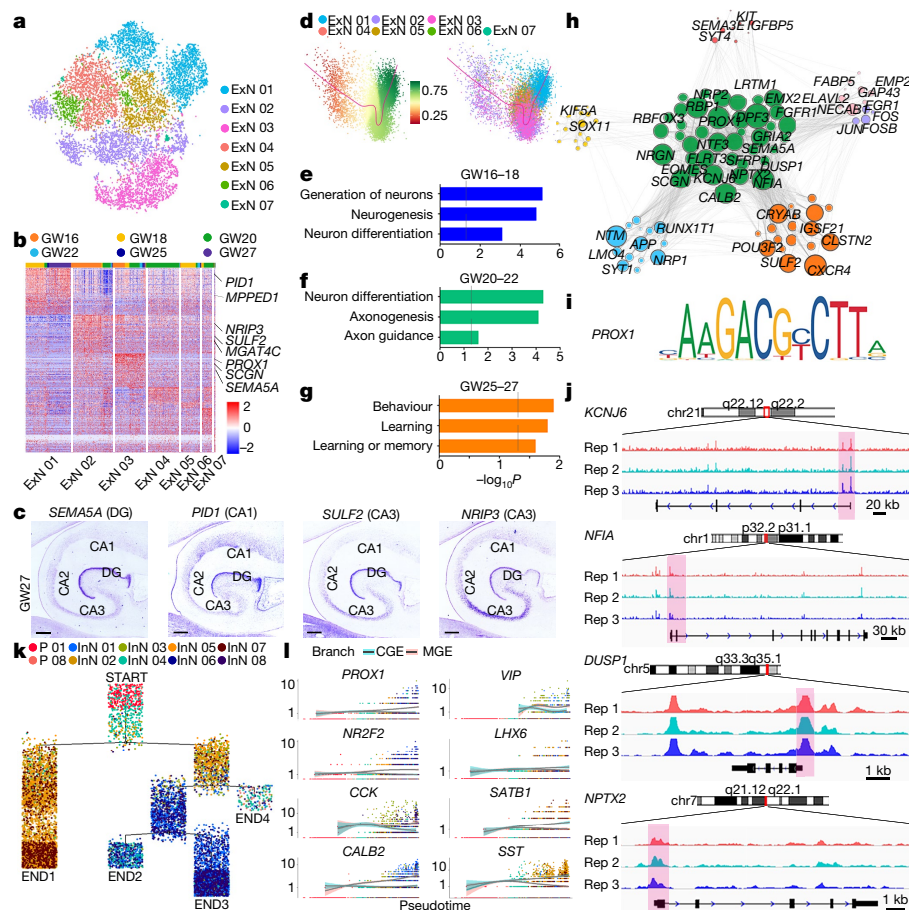
## Neurons in developing hippocampus

To further investigate the developmental characteristics of hippocampal neurons, we subclassified all the excitatory neurons into seven groups by PCA (Fig. 3a, b). Excitatory neurons from CA1, CA3 and DG

were grouped as ExN01–03 (Fig. 3b, Extended Data Fig. 6a). *SEMA5A* and *PID1* were selected as marker genes for DG and CA1, respectively, while *SULF2* and *NR1P3* were considered as CA3 markers (Fig. 3c, Extended Data Fig. 6b). Consistent with progenitor migration paths, the maturation analysis suggests that CA1 neurons were more mature than CA3 and DG neurons (Fig. 3d). Excitatory neurons were categorized into three groups according to their developmental stage, and GO analysis of DEGs indicates that neurogenesis is the major event at GW16–18, followed by axonogenesis (GW20–22) and function development (GW25–27) (Fig. 3e–g). To further analyse the transcriptional regulation of DG formation, we selected the subclusters of highly variable genes and clustered them into nine modules by weighted gene coexpression network analysis (WGCNA) (Extended Data Fig. 6c, d). The green module includes *PROX1*, suggesting that the genes in this module may be correlated with DG development (Fig. 3h). When we analysed ATAC-seq data for the hippocampus at GW25, we found *PROX1* motifs in ATAC peaks close to the TSSs of several genes, including *KCNJ6*, *NFIA*, *DUSP1* and *NPTX2*, which are also in the green module (Fig. 3i, j). Among these genes, *KCNJ6* (also known as *GIRK2*) encodes a member of the G-protein-activated inwardly rectifying K<sup>+</sup> channels that is widely abundant in the brain and has been implicated in learning and memory, reward, motor coordination, and other functions<sup>13</sup>.

Hippocampal inhibitory neurons arise from MGE and CGE precursors. Notably, monocle analysis suggested that the majority of MGE-derived InN (*LHX6*<sup>+</sup>) and CGE-derived InN (*NR2F1/2*<sup>+</sup>) were separated (Fig. 3k, Extended Data Fig. 7a–c). The pseudo-time analysis demonstrated that InN expressing *CCK*, *CALB2* and *VIP* accumulated in the CGE differentiation path, and the majority of *SATB1*<sup>+</sup> and *SS7*<sup>+</sup> neurons were in the MGE path (Fig. 3l, Extended Data Fig. 7b, c). Additionally,





**Fig. 3 | Dynamics of neurogenesis in the developing human hippocampus.**

**a**, Visualization of seven subtypes of excitatory neuron in the developing human hippocampus using *t*-SNE. Sample sizes of clusters: 2,573, 2,347, 1,838, 1,956, 1,192, 717, 92 cells. **b**, Heat map showing the expression level and identity of genes in the excitatory neurons subclasses. Top, distribution of each subclass by gestational week. **c**, In situ hybridization of region-specific genes in DG, CA1 and CA3 at GW27. Scale bar, 600  $\mu$ m. The experiment was repeated three times independently with similar results. **d**, Maturation scores of seven subtypes of excitatory neuron show that CA1 neurons are more mature than CA3 and DG neurons. **e–g**, The enriched gene ontology terms show the cell properties of the hippocampus at different weeks. Sample sizes: 4,912 cells (**e**);

4,164 cells (**f**); 1,639 cells (**g**). **h**, The social network Cytoscape graph depicts the gene network regulation of excitatory neurons. **i, j**, Motifs of *PROX1* (**i**) and the normalized ATAC-seq profile of downstream genes of *PROX1* (**j**) in GW25 hippocampus with three independent biological replicates. **k**, Cell lineage relationships of progenitors and inhibitory neurons analysed in developing human hippocampus. Monocle recovered a branched single-cell trajectory beginning with progenitors and terminating at subgroups of inhibitory neurons. **l**, Markers were ordered by Monocle analysis in pseudo-time. Line with blue shading represents inhibitory neurons derived from CGE; pink shading represents inhibitory neurons derived from MGE.

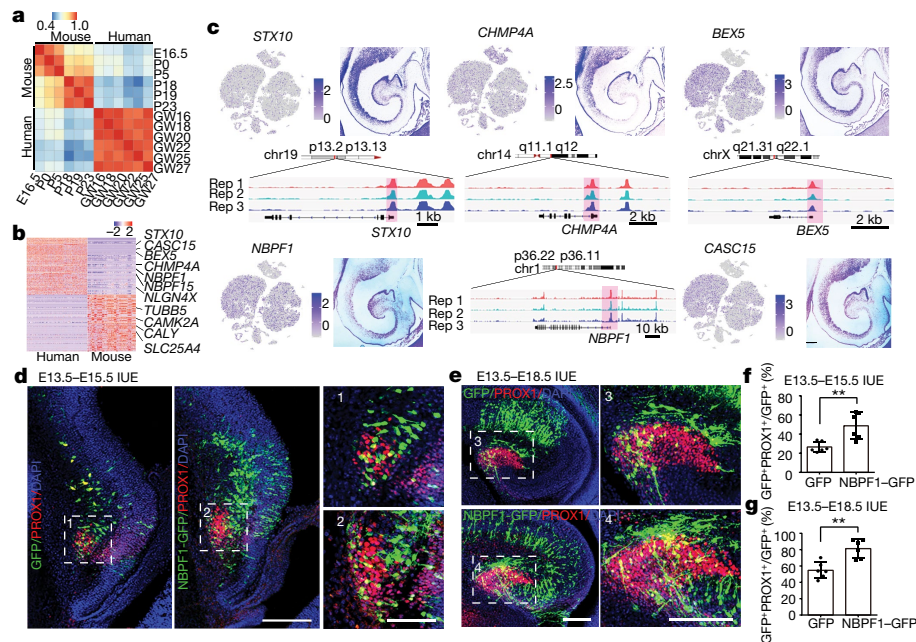
we found genes that may regulate cell fate determination at the first branch point (Extended Data Fig. 7d, e). Microglia, the immune cells in the CNS, originate from the mesoderm<sup>14</sup>. We classified microglia into 11 subclusters and observed that M9 contained microglia in active cell cycles from all developing stages (Extended Data Fig. 8a–d). The immunostaining images also indicated proliferating microglia at GW25 (Extended Data Fig. 8e).

### Evolution signatures of developing hippocampus

Although the hippocampus is considered an evolutionarily conserved part of the brain, transcriptomic correlation coefficient analysis illustrated that the developmental timing of the human hippocampus from GW16 to 20 was similar to that at P0–5 in mice<sup>15,16</sup> (Fig. 4a), suggesting that the human embryonic hippocampal development occurs earlier but lasts for longer than in mice. We also found DEGs in the human hippocampus, some of which are primate-specific, including *STX10*, *CHMP4A*, *BEX5*, *NBPF1* and the long non-coding RNA *CASC15* (Fig. 4b, c). In situ images and ATAC-seq data identified the mRNA localization and transcription regulatory sites of these genes (Fig. 4c). Genes of the neuroblastoma breakpoint family (NBPF) contain a repeated

domain called DUF1220, the copy number of which is related to brain evolution and complexity<sup>17</sup>. Several NBPF family genes are expressed in hippocampal cells, and the expression of *NBPF1* was relatively high and general in all cell types (Fig. 4c, Extended Data Fig. 9a). *NBPF1* with eight DUF1220 domains exists only in primates, and in particular in species that are evolutionally close to humans (Extended Data Fig. 9b, c). To further investigate its role in hippocampal development, we transiently expressed *NBPF1* in the mouse primordial hippocampal area at embryonic day 13.5 (E13.5) and observed that these mice had more *PROX1*<sup>+</sup> cells and an enlarged *PROX1*<sup>+</sup> area at E15.5 and E18.5 when compared with control mice (Fig. 4d–g, Extended Data Fig. 9d–f). To understand how *NBPF1* regulates hippocampal development, we collected single GFP<sup>+</sup> cells (Extended Data Fig. 9g). *LHX2* has been considered as an essential gene in the hippocampal primordium to regulate hippocampal neuronal development<sup>18</sup>. Single-cell quantitative RT–PCR results indicated that *LHX2* expression was higher in *NBPF1*–GFP<sup>+</sup> cells (Extended Data Fig. 9h). Further analysis of open chromatin areas close to the *PROX1* TSS revealed three potential sites for *LHX2* binding (Extended Data Fig. 9i), indicating a possible molecular mechanism by which NBPF1 may regulate hippocampal development via *LHX2*.





**Fig. 4 | Specific genes expressed in the human developing hippocampus.** **a**, Heat map showing correlation of different stages of hippocampus development in human and mouse. The developing human hippocampus is similar to the developing mouse hippocampus at P0–5. **b**, Heat map showing DEGs in human and mouse hippocampus. **c**, Expression of human-specific genes in *t*-SNE plots (left). Right, in situ hybridization at GW25; bottom, normalized ATAC-seq profile of human-specific genes in hippocampus in GW25

## Discussion

We have systematically analysed scRNA-seq and ATAC-seq data to identify cell type diversities, gene expression trajectories, transcription regulation networks and signal transduction pathways in the developing human hippocampus. The hippocampus starts to form from the hippocampal primordium in response to bone morphogenetic protein (BMP) and WNT secreted by the CH<sup>18–20</sup>. An open chromatin area close to the PROX1 TSS contains the binding motif for LEF1 and TCF4, two transcription factors that are involved in the WNT signalling pathway by recruiting the coactivator beta-catenin to enhancer elements of targeting genes<sup>5</sup>, indicating that WNT signals not only initiate differentiation of the medial pallium to the hippocampus, but also contribute to subregional patterning of the hippocampus. The adult neural stem cells located in the subgranular zone give rise to granule cells throughout adult life in most mammals<sup>21</sup>. WNT signalling also helps to regulate granule cell genesis and neural activity in adult mammals<sup>22,23</sup>, indicating that the key gene regulation may be conserved in embryonic and adult neurogenesis in the hippocampal DG.

*HOPX* has been recently identified as a gene that is expressed by dentate precursors and contributes to embryonic and postnatal neurogenesis in mice<sup>24</sup>. Another unbiased single-cell RNA-seq analysis has indicated that perinatal, postnatal, and adult neurogenesis in the mouse DG are fundamentally similar<sup>15</sup>. Notably, clonal lineage-tracing of *HOPX*<sup>+</sup> cells in mice showed that these precursors generate neurons located in the DG or CA<sup>24</sup>. Consistently, we found that at GW11, although most *HOPX*<sup>+</sup> progenitors are located in the DNE, a subset of *HOPX*<sup>+</sup> progenitors is found in the ANE, indicating that *HOPX*<sup>+</sup> progenitors in different locations may have different cell fates.

The copy number of the DUF1220 protein domain in the genome is correlated with the evolutionary proximity of the species to humans as well as with brain size, cognitive capability, and severity of autism<sup>17,25–27</sup>. Major copies of human DUF1220 domains are encoded by the *NBPF* gene family. Microarray data from the Allen Brain Atlas suggest that *NBPF*1

with three independent biological replicates. Scale bar, 300  $\mu$ m.

**d, e**, Overexpression of *NBPF*1 promotes DG formation at E13.5, observed at E15.5 (**d**) and E18.5 (**e**) in mouse. Scale bars, 500  $\mu$ m (**d**, left), 100  $\mu$ m (**d**, right), 200  $\mu$ m (**e**). **f, g**, Percentage of PROX1<sup>+</sup> cells among GFP<sup>+</sup> cells. **f**, E13.5–E15.5:  $**P = 0.0049$ , two-sided *t*-test.  $n = 6$ , 5 brain slices per experiment; mean  $\pm$  s.d. **g**, E13.5–E18.5:  $**P = 0.0015$ .  $n = 6$ , 5 brain slices per experiment. IUE, in utero electroporation.

expression decreases when the human brain develops (<http://www.brainspan.org>). *LHX2* is expressed in the dorsal and medial pallium but not in the CH, which secretes WNT ligands and functions as an organizer that is necessary and sufficient to induce the hippocampus<sup>18</sup>. Notably, expression of *NBPF*1 upregulates *LHX2* expression and increases the number of hippocampal PROX1<sup>+</sup> granule cells in the developing mouse brain. However, the detailed molecular mechanisms of this process need further investigation.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-019-1917-5>.

- Bird, C. M. & Burgess, N. The hippocampus and memory: insights from spatial processing. *Nat. Rev. Neurosci.* **9**, 182–194 (2008).
- Miller, A. M., Vedder, L. C., Law, L. M. & Smith, D. M. Cues, context, and long-term memory: the role of the retrosplenial cortex in spatial cognition. *Front. Hum. Neurosci.* **8**, 586 (2014).
- Lavado, A., Lagutin, O. V., Chow, L. M., Baker, S. J. & Oliver, G. Prox1 is required for granule cell maturation and intermediate progenitor maintenance during brain neurogenesis. *PLoS Biol.* **8**, e1000460 (2010).
- Sugiyama, T., Osumi, N. & Katsuyama, Y. The germinal matrices in the developing dentate gyrus are composed of neuronal progenitors at distinct differentiation stages. *Dev. Dyn.* **242**, 1442–1453 (2013).
- Galceran, J., Miyashita-Lin, E. M., Devaney, E., Rubenstein, J. L. R. & Grosschedl, R. Hippocampus development and generation of dentate gyrus granule cells is regulated by LEF1. *Development* **127**, 469–482 (2000).
- Lee, S. M. K., Tole, S., Grove, E. & McMahon, A. P. A local Wnt-3a signal is required for development of the mammalian hippocampus. *Development* **127**, 457–467 (2000).
- Zhong, S. J. et al. A single-cell RNA-seq survey of the developmental landscape of the human prefrontal cortex. *Nature* **555**, 524–528 (2018).
- Nowakowski, T. J. et al. Spatiotemporal gene expression trajectories reveal developmental hierarchies of the human cortex. *Science* **358**, 1318–1323 (2017).

9. Mu, L. et al. SoxC transcription factors are required for neuronal differentiation in adult hippocampal neurogenesis. *J. Neurosci.* **32**, 3067–3080 (2012).
10. Wang, Y., Lin, L., Lai, H., Parada, L. F. & Lei, L. Transcription factor Sox11 is essential for both embryonic and adult neurogenesis. *Dev. Dyn.* **242**, 638–653 (2013).
11. Bandler, R. C., Mayer, C. & Fishell, G. Cortical interneuron specification: the juncture of genes, time and geometry. *Curr. Opin. Neurobiol.* **42**, 17–24 (2017).
12. Pollen, A. A. et al. Molecular identity of human outer radial glia during cortical development. *Cell* **163**, 55–67 (2015).
13. Cooper, A. et al. Trisomy of the G protein-coupled K<sup>+</sup> channel gene, *Kcnj6*, affects reward mechanisms, cognitive functions, and synaptic plasticity in mice. *Proc. Natl Acad. Sci. USA* **109**, 2642–2647 (2012).
14. Gomez Perdiguer, E., Schulz, C. & Geissmann, F. Development and homeostasis of “resident” myeloid cells: the case of the microglia. *Glia* **61**, 112–120 (2013).
15. Hochgerner, H., Zeisel, A., Lönnerberg, P. & Linnarsson, S. Conserved properties of dentate gyrus neurogenesis across postnatal development revealed by single-cell RNA sequencing. *Nat. Neurosci.* **21**, 290–299 (2018).
16. La Manno, G. et al. RNA velocity of single cells. *Nature* **560**, 494–498 (2018).
17. Dumas, L. J. et al. DUF1220-domain copy number implicated in human brain-size pathology and evolution. *Am. J. Hum. Genet.* **91**, 444–454 (2012).
18. Mangale, V. S. et al. Lhx2 selector activity specifies cortical identity and suppresses hippocampal organizer fate. *Science* **319**, 304–309 (2008).
19. Grove, E. A., Tole, S., Limon, J., Yip, L. & Ragsdale, C. W. The hem of the embryonic cerebral cortex is defined by the expression of multiple Wnt genes and is compromised in Gli3-deficient mice. *Development* **125**, 2315–2325 (1998).
20. Yoshida, M., Assimacopoulos, S., Jones, K. R. & Grove, E. A. Massive loss of Cajal-Retzius cells does not disrupt neocortical layer order. *Development* **133**, 537–545 (2006).
21. Kempermann, G., Song, H. & Gage, F. H. Neurogenesis in the adult hippocampus. *Cold Spring Harb. Perspect. Biol.* **7**, a018812 (2015).
22. Ortiz-Matamoros, A., Salcedo-Tello, P., Avila-Muñoz, E., Zepeda, A. & Arias, C. Role of wnt signaling in the control of adult hippocampal functioning in health and disease: therapeutic implications. *Curr. Neuropharmacol.* **11**, 465–476 (2013).
23. Varela-Nallar, L. & Inestrosa, N. C. Wnt signaling in the regulation of adult hippocampal neurogenesis. *Front. Cell. Neurosci.* **7**, 100 (2013).
24. Berg, D. A. et al. A common embryonic origin of stem cells drives developmental and adult neurogenesis. *Cell* **177**, 654–668.e15 (2019).
25. Davis, J. M. et al. DUF1220 dosage is linearly associated with increasing severity of the three primary symptoms of autism. *PLOS Genet.* **10**, e1004241 (2014).
26. Keeney, J. G., Dumas, L. & Sikela, J. M. The case for DUF1220 domain dosage as a primary contributor to anthropoid brain expansion. *Front. Hum. Neurosci.* **8**, 427 (2014).
27. Popesco, M. C. et al. Human lineage-specific amplification, selection, and neuronal expression of DUF1220 domains. *Science* **313**, 1304–1307 (2006).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2020

## Methods

No statistical methods were used to predetermine sample size. The experiments were not randomized and investigators were not blinded to allocation during experiments and outcome assessment.

### Tissue sample collection

The de-identified human tissue collection and research protocols were approved by the Reproductive Study Ethics Committee of Beijing Anzhen Hospital and the institutional review board (ethics committee) of the Institute of Biophysics. The informed consent was designed as recommended by the ISSCR guidelines for fetal tissue donation and fetal tissue samples were collected after the donor patients signing an informed consent document that was in strict observance of the legal and institutional ethical regulations for samples from elective pregnancy terminations at Beijing Anzhen Hospital, Capital Medical University. All samples used in these studies had not been involved in any other procedures. All the protocols were in compliance with the Interim Measures for the Administration of Human Genetic Resources, administered by the Ministry of Science and Technology of China.

### Animals

Timed pregnant female mice at embryonic day 13.5 were used for in utero electroporation experiments. Embryos for experiments after in utero electroporation included both male and female mice. Mouse housing and experimental protocols in this study were in compliance with the guidelines of the Institutional Animal Care and Use Committee of the Institute of Biophysics, CAS. All mice had free access to food and water and were housed in the institutional animal care facility with a 12-h light–dark schedule.

### Tissue sample dissection

Gestational age was measured in weeks from the first day of the woman's last menstrual cycle to the sample collecting date. Fetal brains were collected in ice-cold artificial cerebrospinal fluid containing 125.0 mM NaCl, 26.0 mM NaHCO<sub>3</sub>, 2.5 mM KCl, 2.0 mM CaCl<sub>2</sub>, 1.0 mM MgCl<sub>2</sub>, 1.25 mM NaH<sub>2</sub>PO<sub>4</sub> at a pH of 7.4 when oxygenated (95% O<sub>2</sub> and 5% CO<sub>2</sub>). The hippocampus was dissected and put in Hibernate E medium (Invitrogen, Cat. A1247601). The hippocampus tissue was first digested in 2 mg/ml collagenase IV (Gibco, Cat. 17104-019) and 10 U/μl DNase I (NEB, Cat. M0303L) in Hibernate E medium and then in 1 mg/ml papain (Sigma, Cat. P4762) and 10 U/μl DNase I in Hibernate E medium. Samples were vortexed at 300g and 37 °C on a thermocycler for 20 min. Further pipetting was used to fully digest the tissue into single cells. After that, the cell suspension was centrifuged at 700g for 5 min to obtain the cell pellet. The digestion medium was carefully removed and the cell pellet was resuspended in 300 μl 0.04% BSA in PBS and kept on ice.

### RNA library preparation for high-throughput sequencing

Thousands of cells were partitioned into nanolitre-scale Gel Bead-In-EMulsions (GEMs) using 10x GemCode Technology, where cDNA produced from the same cell shares a common 10x Barcode. Upon dissolution of the single cell 3' gel bead in a GEM, primers containing an Illumina R1 sequence (read1 sequencing primer), a 16-bp 10x Barcode, a 10-bp randomer and a poly-dT primer sequence were released and mixed with cell lysate and Master Mix. After incubation of the GEMs, barcoded, full-length cDNA from poly-adenylated mRNA was generated. Then the GEMs were broken and silane magnetic beads were used to remove leftover biochemical reagents and primers. Prior to library construction, enzymatic fragmentation and size selection were used to optimize the cDNA amplicon size. P5, P7, a sample index and R2 (read 2 primer sequence) were added to each selected cDNA during end repair and adaptor ligation. P5 and P7 primers were used in Illumina bridge amplification of the cDNA (<http://10xgenomics.com>). Finally, the library was sequenced into 150-bp paired-end reads using the Illumina HiSeq4000.

### Data processing of scRNA-seq from Chromium system

Cell Ranger 2.0.1 (<http://10xgenomics.com>) was used to perform quality control and read counting of Ensemble genes with default parameters (v2.0.1) by mapping to the hg19 human genome. We excluded poor-quality cells after the gene-cell data matrix was generated by Cell Ranger software using the Seurat package (v2.3.4). Only cells that expressed more than 800 genes and fewer than 7,000 genes were considered, and only genes expressed in at least 30 single cells (0.1% of the raw data) were included for further analysis. Cells that expressed haemoglobin genes (*HBM*, *HBA1*, *HBA2*, *HBB*, *HBD*, *HBE1*, *HBG1*, *HBG2*, *HBQ1* and *HBZ*) were also excluded. Cells with a mitochondrial gene percentage over 15% were discarded. In total, 17,737 genes across 30,416 single cells remained for subsequent analysis. The data were normalized to a total of  $1 \times 10^4$  molecules per cell for the sequencing depth using the Seurat package. The batch effect was mitigated by using the ScaleData function of Seurat (v2.3.4).

### Identification of cell types and subtypes by dimensional reduction and PAGA analysis

The Seurat package (v2.3.4) was used to perform linear dimensional reduction. We selected 982 highly variable genes with average expression between 0.0125 and 8 and dispersion greater than 2 as input for PCA. Then we identified significant PCs based on the JackStrawPlot function. Strong PC1–PC10 were used for *t*-SNE to cluster the cells by FindClusters function with resolution 1.2. Clusters were identified by the expression of known cell-type markers and GO analysis. The markers *ASCL1*, *NEUROD2*, *GAD1*, *OLIG2*, *MBP*, *AQP4*, *SPARC* and *PTPRC* were used to hippocampal cells as progenitor cells, excitatory neurons, inhibitory neurons, OPCs, oligodendrocytes, astrocytes, endothelial cells and microglia, respectively.

Three-dimensional *t*-SNE was applied to cluster all cells in the human developing hippocampus (dim.embed = 5) with PC1–PC10. Visualizations were done using *rgl* package (v0.99.16) implemented in R. We then applied partition-based graph abstraction (PAGA) to predict a lineage tree for the hippocampal and the prefrontal cortical cells in an unbiased way. We produced a consolidated lineage tree that included all identified cell types rooted to a stem cell group.

### Identification of DEGs among clusters

The DEGs of each cluster were identified using the FindAllMarkers function (thresh.use = 0.25, test.use = “wilcox”) with the Seurat R package (6). We used the Wilcoxon rank-sum test (default), and genes with average expression difference >0.5 natural log with *P* < 0.05 were selected as marker genes. Enriched GO terms of marker genes were identified using DAVID 6.8<sup>28,29</sup> (<https://david.ncifcrf.gov/home.jsp>) and Metascape<sup>30</sup> (<http://metascape.org>).

### Constructing single cell trajectories in the hippocampus

The Monocle 2 R package (version 2.6.4) and Monocle 3 alpha R package (version 2.99.2) were applied to construct single cell pseudo-time trajectories to discover developmental transitions<sup>31–33</sup>. We used highly variable genes identified by Seurat to sort cells into pseudo-time order. The actual gestational time of each cell informs us which states of cells are at the beginning of pseudo-time in the first round of “orderCells”. We then call “orderCells” again, passing this state as the root\_state argument. “DDRTree” and “UMAP” were applied to reduce dimensional space and the minimum spanning tree on cells was plotted using the visualization functions “plot\_complex\_cell\_trajectory” or “plot\_3d\_cell\_trajectory” for Monocle 2 and Monocle 3 alpha, respectively.

### Cell-cycle analysis

In the cell-cycle analysis, we applied a cell-cycle related gene set with 43 genes expressed during G1/S and 54 genes expressed during G2/M<sup>34,35</sup>. We defined the G1/S and G2/M states of each cell by comparing

# Article

the average expression of the two gene sets using the CellCycleScoring function using Seurat R package. These gene sets should be anticorrelated in their expression levels, and cells expressing neither are likely to be in the G1 phase (not cycling).

## WGCNA analysis in categorizing genes

WGCNA analysis was performed by R package “WGCNA”<sup>36,37</sup> (R version 3.4.3, <https://cran.r-project.org/src/contrib/Archive/WGCNA>; package version 1.6.6). The WGCNA soft power value was determined by navigating the soft-threshold-mean-connectivity curve. Modules with <0.25 similarity were merged. Modules correlated with a specific cell subtype were considered as standard modules for categorizing genes into certain cell subtypes. Seven modules were selected for neuron subtypes.

## ATAC library preparation for high-throughput sequencing

ATAC-seq was performed as described previously<sup>38,39</sup>. In brief, a total of 50,000 cells were washed twice with 50 µl of cold PBS and resuspended in 50 µl lysis buffer (10 mM Tris-HCl pH 7.4, 10 mM NaCl, 3 mM MgCl<sub>2</sub>, 0.1% (v/v) Nonidet P40 Substitute). The suspension of nuclei was then centrifuged for 10 min at 500g at 4 °C, followed by the addition of 50 µl transposition reaction mix (10 µl 5 × TTBL buffer, 4 µl TTE mix and 36 µl nuclease-free H<sub>2</sub>O) from the TruePrep DNA Library Prep Kit V2 for Illumina (Vazyme Biotech). Samples were then incubated at 37 °C for 30 min. DNA was isolated using a QIAquick PCR Purification Kit (QIAGEN). ATAC-seq libraries were first subjected to five cycles of pre-amplification. To determine the suitable number of cycles required for the second round of PCR, the library was assessed by quantitative PCR as described previously<sup>38</sup> and then PCR amplified for the appropriate number of cycles. Libraries were purified with a QIAquick PCR Purification Kit (QIAGEN). Library quality was checked using a High Sensitivity DNA Analysis Kit (Agilent). Finally, 2 × 150 paired-end sequencing was performed on an Illumina HiSeq X-10.

## ATAC-seq data analysis

In simple terms, we removed adaptor sequences and then mapped reads to the hg19 reference genome with the parameters: -t -q -N1 -L 25 -X 2000 using Bowtie2 (version 2.3.4.3). All unmapped reads, non-uniquely mapped reads and PCR duplicates were removed. The uniquely mapped reads were shifted by +4 or -5 bp according to the strand of the read. To visualize the ATAC-seq signal, we extended each read by 50 bp and counted the coverage for each base. All the ATAC-seq peaks were called by MACS2 v2.1.1 with the parameters -nolambda.

## ATAC-seq data quality control

ATAC-seq data quality was evaluated for several parameters, including the number of raw reads, alignment rate, percentage of reads mapped to chromosome M, percentage of reads mapped to repeat regions (black list), percentage of reads that passed MAPQ score filter, percentage of total signal within known artefact regions and correlation between replications.

## Connecting transcription factors to target genes

To find the potential transcription factors that bind the *PROX1* regulatory sequence (TSS ± 2k), FIMO from MEME Suite (version 5.0.4) was used for motif enrichment analysis. To investigate the genes that are regulated by *PROX1*, the *PROX1* motif profile was downloaded from the Jasp database (<http://jaspar.genereg.net/>), and we used FIMO from the MEME suite for enrichment analysis of our peaks.

## Immunofluorescent staining

Tissue samples were fixed overnight in 4% paraformaldehyde, cryoprotected in 30% sucrose, and embedded in optimal cutting temperature (Thermo Scientific). Thin 40-µm cryosections were collected on superfrost slides (VWR) using a Leica CM3050S cryostat. For immunohistochemistry, heat-induced antigen retrieval was performed in 10

mM sodium citrate buffer, pH 6. Primary antibodies: mouse anti-CD45 (1:100, Abcam ab8216), goat anti-SOX2 (1:250, Santa Cruz sc-17320), rabbit anti-PAX6 (1:500, BioLegend 901301), rabbit anti-NEUROD2 (1:500, Abcam ab104430), mouse anti-NEUROD1 (1:100, Abcam ab60704), rabbit anti-HOPX (1:1,000, Santa Cruz sc-30216), mouse anti-Ki67 (1:100, BD 550609), mouse anti-SATB2 (1:250, Abcam ab51502), mouse anti-MEIS2 (1:200, Santa Cruz sc-81986), rabbit anti-PROX1 (1:500, Abcam ab199359), rabbit anti-OLIG2 (1:500, Millipore AB9610), human anti-MBP (1:1,000, Abcam ab209328), mouse anti-GFAP (1:200, CST 3670S) diluted in blocking buffer containing 10% donkey serum, 0.5% Triton-X100 and 0.2% gelatin diluted in PBS at pH 7.4. Binding was revealed using an appropriate Alexa Fluor 488, Alexa Fluor 594, or Alexa Fluor 647 fluorophore-conjugated secondary antibody (Life Technologies). Cell nuclei were counterstained using DAPI (Life Technologies). Images were collected using an Olympus FV1000 confocal microscope.

## In situ hybridization

The in situ hybridization protocol has been described previously<sup>40</sup>. In brief, probes complementary to target human mRNA used for RNA in situ hybridization were cloned from primary human fetal cortical cDNA samples and reverse-transcribed using PrimeScript II 1st Strand cDNA Synthesis Kit (Takara) with oligo dT primers. Total RNA was isolated from GW27 human hippocampus using SV Total RNA Isolation System (Promega). Specific genes were amplified using the following primers: *SEMASA* forward AGC TCG CTT GGC TTT AGT CTT A, reverse CAA AAT AGG CTT TGA CTC CCA C; *PID1* forward TGG GAT CTC TAG TGG GGT GG, reverse TAA GGC TTC TTA GGT GCC GC; *SULF2* forward GTT TGA CAT CAG GGT CCC GT, reverse CTT TAA TGG GGT TGG CGG CT; *NRIP3* forward AGC TGT GGT TGA TGA CAA TGA G, reverse CTG TAA TGG ATA ATG TCC CTG G; *STX10* forward GGG GAA GGG ACT GAC ATG TC, reverse GGA GGG CTG GGG TCA GAG AG; *CHMP4A* forward GAT TGG GCA AGG CTG GTC CC, reverse TTG GGA GCT GGC CCT GCC GG; *BEX5* forward TCA ACA TGG AAA ATG TCC CC, reverse AGA CTG CTT TTA AAT TGC TT; *NBPFL* forward GGG TGC ACC AAG AGC AGC CT, reverse CCT CAG CAT AAA TTT TAT GA; *CASC15* forward CAA GCA TGT AGC CCT GCC CG, reverse CTC TGT TTC TGT CAT CTC TC; primers specific to target genes of interest were designed using Primer3 and amplified by PCR using Q5 High-Fidelity DNA Polymerase (NEB). PCR products of predicted band size were gel extracted and ligated into the Hieff Clone Plus One Step Cloning Kit (Yeason). Ligation products were transfected into Trans5α Chemically Competent *E. coli* (Transgene). Cloned sequences were confirmed by sequencing. Digoxigenin-labelled RNA probes for in situ hybridization were generated by linearizing the pSPT18 Vector and in vitro transcribing the probe using T7 or SP6 RNA Polymerase (Roche) in the presence of DIG-RNA Labelling Mix (Roche). Fetal brain sections of 30µm thickness were hybridized with RNA probes at a final concentration of 500 ng/ml overnight at 64.5 °C in hybridization solution (50% formamide, 10% dextran sulfate, 0.2% tRNA (Invitrogen), 1 × Denhardt's solution (Sigma) and 1 × salt solution (containing 0.2 M NaCl, 0.01 M Tris, 5 mM NaH<sub>2</sub>PO<sub>4</sub>, 5 mM Na<sub>2</sub>HPO<sub>4</sub>, 5 mM EDTA pH 7.5)) overnight. After the sections were washed, alkaline phosphatase-coupled anti-digoxigenin Fab fragments (Roche) were applied. For visualization of the labelled cRNAs, the sections were incubated in the dark in NBT/BCIP solution (Roche). Images were taken using a Leica SCN400 (Leica Microsystems).

## Plasmids and in utero electroporation

*NBPFL* genes were cloned into a pEGFP-C1 vector. Electroporation was performed as previously described<sup>41</sup>. In brief, timed pregnant CD-1 mice (E13.5) were deeply anaesthetized with isoflurane, and the uterine horns were exposed through a midline incision. 1 µl of plasmid DNA (1–2 µg/µl) mixed with Fast Green (Sigma) was manually microinjected into the fetal brain lateral ventricle through the uterus, using a bevelled and calibrated glass micropipette (Drummond Scientific) followed by five 50-ms pulses of 50 mV with a 1 s interval delivered across the

uterus with two 9-mm electrode paddles positioned on either side of the head (BTX, ECM830).

### Patch-qRT-PCR of NBPF1-GFP plasmid overexpressed cells

Coronal slices containing cells overexpressing the NBPF1-GFP plasmid were prepared using a vibratome (VT1200S, Leica, Wetzlar, Germany) in oxygenated (95% O<sub>2</sub> and 5% CO<sub>2</sub>) ice-cold sucrose-based artificial cerebrospinal fluid (s-ACSF, 234 mM sucrose, 2.5 mM KCl, 26 mM NaHCO<sub>3</sub>, 1.25 mM NaH<sub>2</sub>PO<sub>4</sub>, 11 mM D-glucose, 0.5 mM CaCl<sub>2</sub> and 10 mM MgSO<sub>4</sub>). The slices were kept in an incubating chamber filled with oxygenated ACSF (126 mM NaCl, 3 mM KCl, 1.2 mM NaH<sub>2</sub>PO<sub>4</sub>, 2.4 mM CaCl<sub>2</sub>, 1.3 mM MgSO<sub>4</sub>, 26 mM NaHCO<sub>3</sub>, 10 mM D-glucose) at 34 °C for 30 min. After a recovery period of at least 60 min at room temperature, an individual slice was transferred to a recording chamber and was continuously superfused with oxygenated ACSF (4 ml/min) at room temperature. We captured whole cells overexpressing the NBPF1-GFP plasmid and distributed each into a single tube, and then we used SMART-seq2 to amplify the mRNA into a cDNA library. Then, we used qRT-PCR to detect *NBPF1* and *LHX2* gene expression. Specific genes were amplified using the following primers: *GAPDH* forward GTC AAG CTC ATT TCC TGG TAT GAC, reverse TAT GGG GGT CTG GGA TGG AA; *NBPF1* forward GCG AGG CTG CCC GAG CTT CT, reverse GAC TTC GCG TAA CTT CCC ATT CA; *LHX2* forward GAA CGA TGC TGA ACA CCT GG, reverse AAC CAG ACC TGG AGG AC TCT C.

### Statistical analysis

Comparisons between two groups were made using *t*-tests. The quantification graphs were analysed by using GraphPad Prism (GraphPad Software). Sample size and *P* values are given in the Figure legends.

### Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

### Data availability

The scRNA-seq data and ATAC-seq data used in this study have been deposited in the Gene Expression Omnibus (GEO) under accession number GSE131258. Raw image files used in the figures that support the findings of this study are available from the corresponding authors upon reasonable request.

28. Huang, W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **37**, 1–13 (2009).

29. Huang, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009).
30. Tripathi, S. et al. Meta- and orthogonal integration of influenza “OMICs” data defines a role for UBR4 in virus budding. *Cell Host Microbe* **18**, 723–735 (2015).
31. Qiu, X. et al. Single-cell mRNA quantification and differential analysis with Census. *Nat. Methods* **14**, 309–315 (2017).
32. Qiu, X. et al. Reversed graph embedding resolves complex single-cell trajectories. *Nat. Methods* **14**, 979–982 (2017).
33. Trapnell, C. et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* **32**, 381–386 (2014).
34. Macosko, E. Z. et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202–1214 (2015).
35. Tirosh, I. et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* **352**, 189–196 (2016).
36. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559 (2008).
37. Langfelder, P. & Horvath, S. Fast R functions for robust correlations and hierarchical clustering. *J. Stat. Softw.* **46**, i11 (2012).
38. Buenostro, J. D., Wu, B., Chang, H. Y. & Greenleaf, W. J. ATAC-seq: A method for assaying chromatin accessibility genome-wide. *Curr. Protoc. Mol. Biol.* **109**, 21–29 (2015).
39. Buenostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* **10**, 1213–1218 (2013).
40. Palop, J. J., Roberson, E. D. & Cobos, I. Step-by-step in situ hybridization method for localizing gene expression changes in the brain. *Methods Mol. Biol.* **670**, 207–230 (2011).
41. Wang, X., Tsai, J. W., LaMonica, B. & Kriegstein, A. R. A new subtype of progenitor cell in the mouse embryonic neocortex. *Nat. Neurosci.* **14**, 555–561 (2011).

**Acknowledgements** This work was supported by the National Key R&D Program of China (2019YFA0110100), the Strategic Priority Research Program of the Chinese Academy of Sciences (XDA16020601, XDB32010100), the National Basic Research Program of China (2017YFA0102601, 2017YFA0103303), the National Natural Science Foundation of China (NSFC) (91732301, 31671072, 31771140, 81891001), the Grants of Shanghai Brain-Intelligence Project from STCSM (16JC1420500), the Grants of Beijing Brain Initiative of Beijing Municipal Science & Technology Commission (Z181100001518004).

**Author contributions** Q.W. and X.W. conceived the project, designed the experiments and wrote the manuscript. S. Zhong and Y.L. performed the scRNA-seq experiment. W.D. performed the ATAC-seq and animal surgery. S. Zhong, Y.L., X.F., S. Zhang and F.T. analysed the RNA-seq data. S. Zhong and H.D. analysed the ATAC-seq data. L.S. and R.C. collected single cells by patch-clamping. Z.L. and S. Zhong carried out qRT-PCR. Q.M. prepared the samples. S. Zhong, Q.W. and W.D. performed immunostaining, in situ hybridization and imaging. All authors edited and proofread the manuscript.

**Competing interests** The authors declare no competing interests.

### Additional information

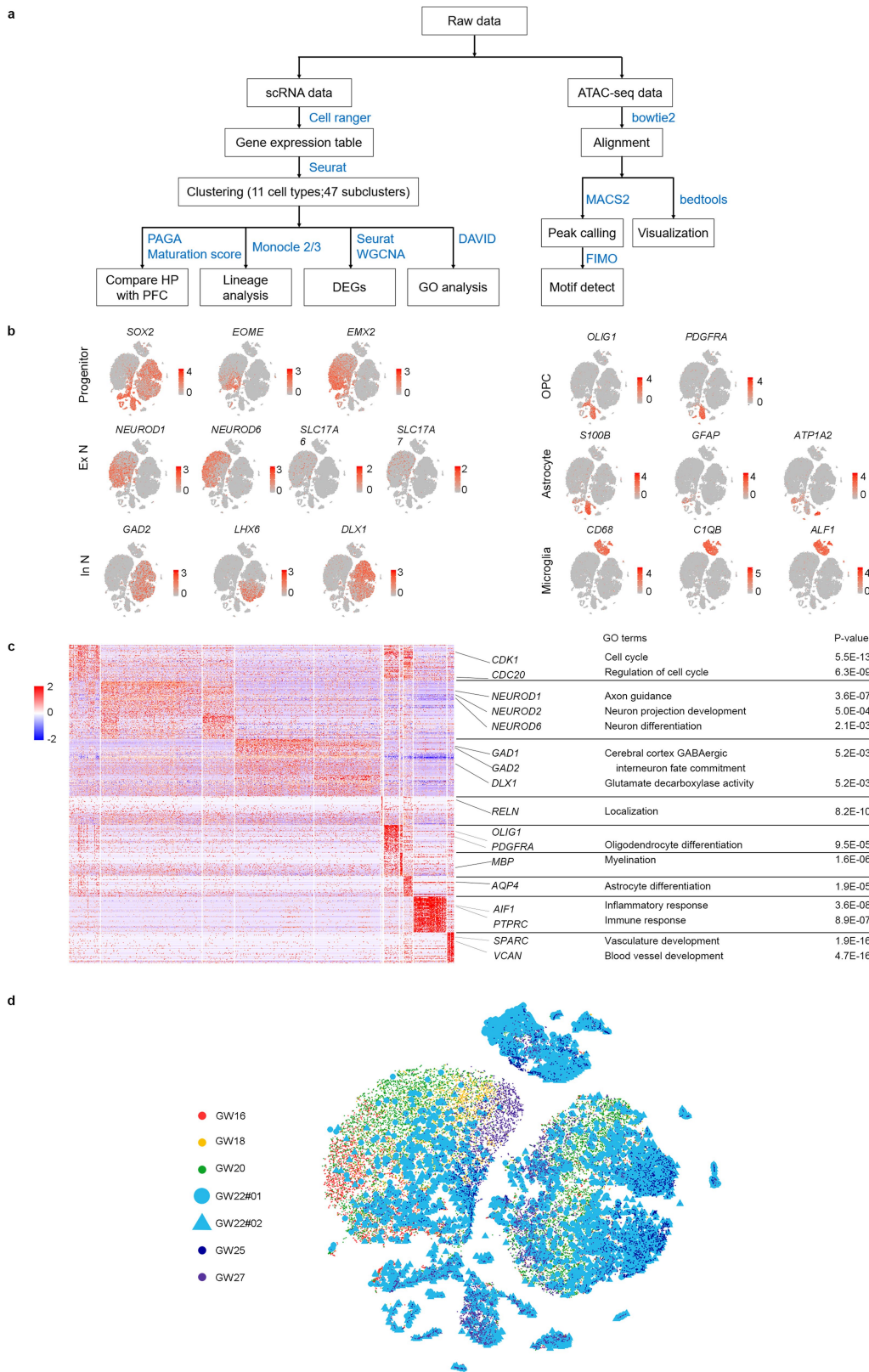
**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41586-019-1917-5>.

**Correspondence and requests for materials** should be addressed to Q.W. or X.W.

**Peer review information** Nature thanks Joseph Loturco, Christopher Walsh and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

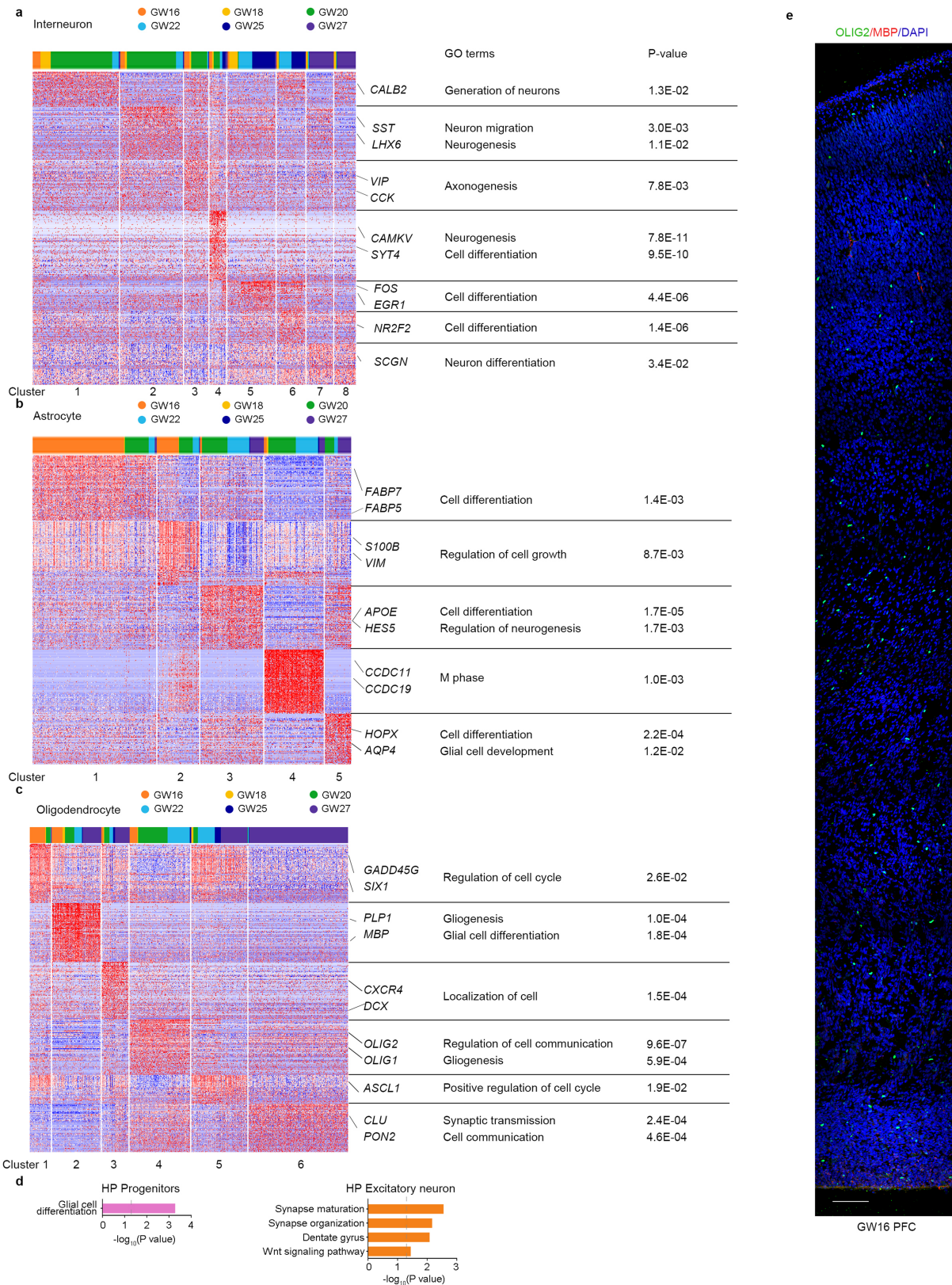
**Reprints and permissions information** is available at <http://www.nature.com/reprints>.





**Extended Data Fig. 1 | Single-cell RNA-seq information and molecular diversity of single cells. a**, Scheme of bioinformatic analysis. **b**, Expression of known markers shown using the same layout as in Fig. 1b. Grey, no expression; red, relative expression. **c**, Heat map showing the expression level and identity of genes in all cells in the developing hippocampus. Sample sizes: astrocytes, 703; Cajal–Retzius cells, 101; endothelial cells, 540; non-DG ExN, 8,199; MGE-derived InN, 6,377; microglia, 2,660; oligodendrocytes, 209; OPCs, 1,250;

progenitors, 2,486; DG ExN, 2,516; CGE-derived InN, 5,375. **d**, *t*-SNE plots of cells in the hippocampus. Two repetitions of GW22 are labelled in different shapes, and no obvious distribution differences are observed among the different batches from the same embryonic stages. Each cell colour represents the gestational week. Sample size: GW16, 4,411 cells; GW18, 4,035 cells; GW20, 10,101 cells; GW22#01, 1,617 cells; GW22#02, 2,485 cells; GW25, 2,824 cells; GW27, 4,943 cells.

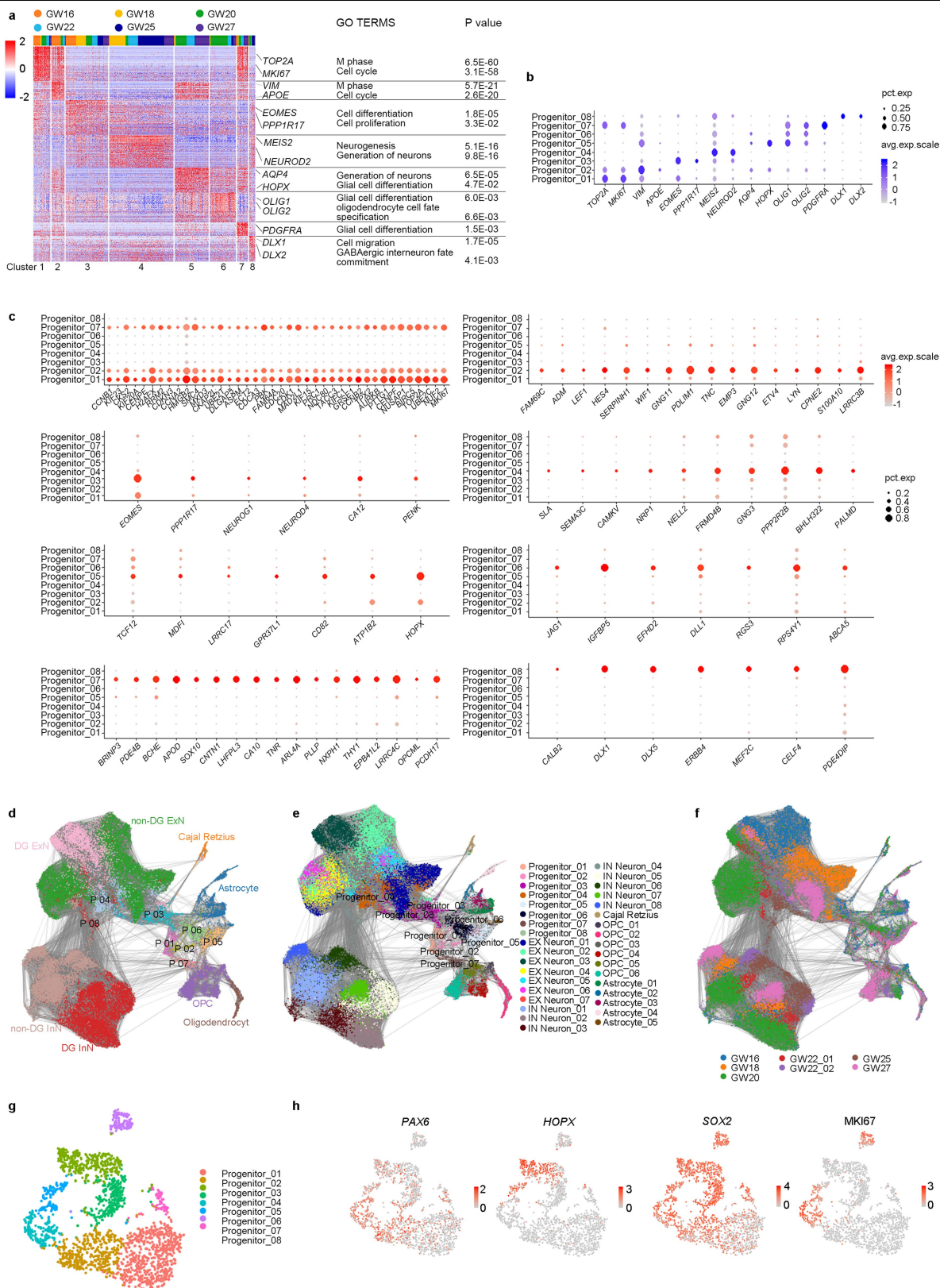


**Extended Data Fig. 2 | Molecular diversity of subgroups of cells.**

**a–c**, Heat maps show the subclasses of inhibitory neurons (**a**), astrocytes (**b**) and oligodendrocytes (**c**). The genes are organized into clusters. The bar chart on the top shows the gestational week. Specific genes related to each subtype are highlighted on the right with enriched GO terms. Interneurons: 3,189, 2,334, 909, 670, 1,765, 1,073, 1,019, 793 cells; astrocytes: 275, 95, 141, 134, 58

cells; oligodendrocytes: 103, 227, 131, 282, 257, 459 cells. **d**, The enriched GO terms show the cell properties of the hippocampus in different cell types. Progenitors, 2,486 cells; excitatory neurons, 10,715 cells. **e**, Immunostaining for oligodendrocyte markers at GW16 showing the position and morphology of oligodendrocytes in human prefrontal cortex. Scale bar, 500  $\mu\text{m}$ . The experiment was repeated three times independently with similar results.

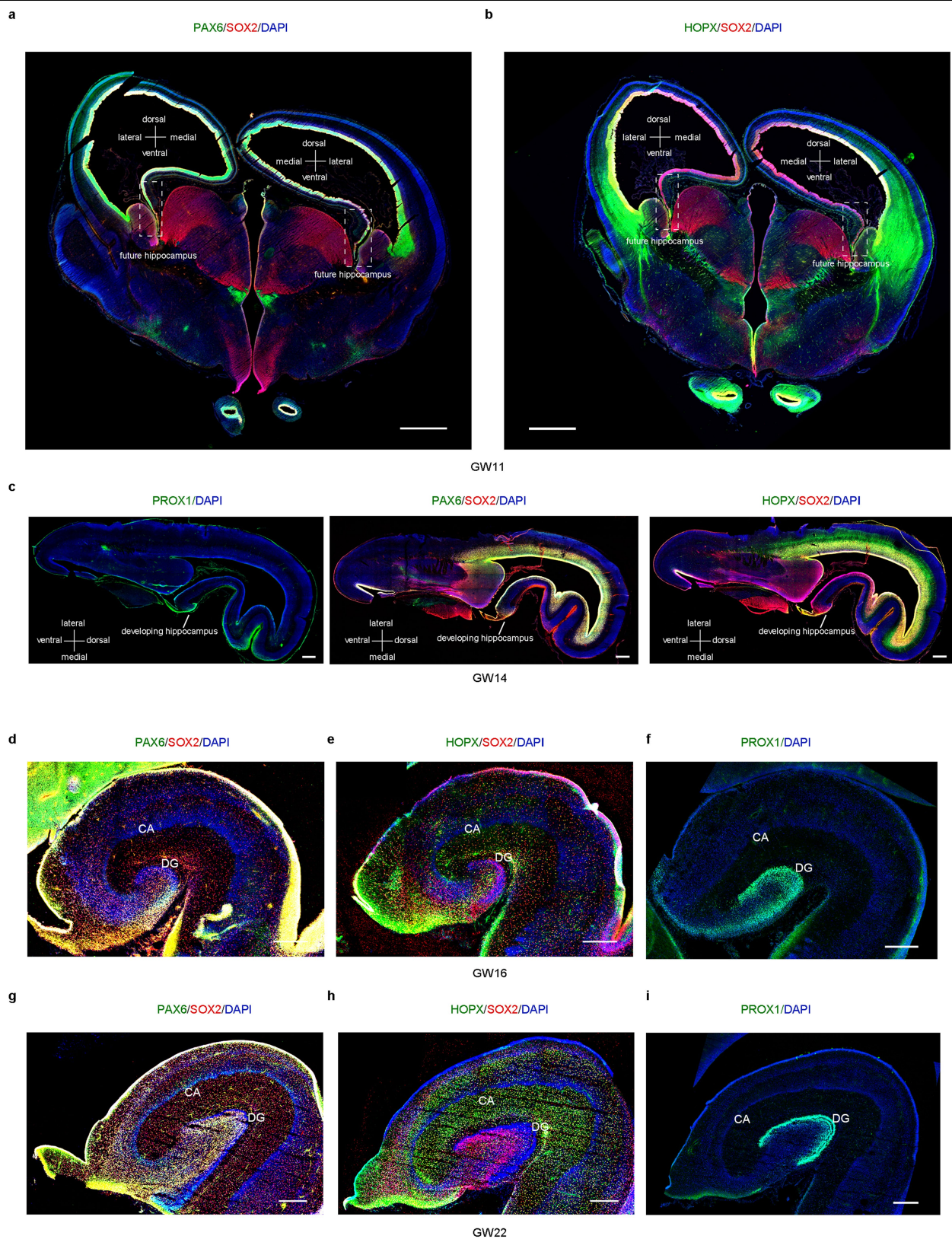




**Extended Data Fig. 3** | See next page for caption.

**Extended Data Fig. 3 | Molecular diversity of progenitors in the hippocampus.** **a**, Heat map showing the expression levels and identities of genes in the progenitor subclasses. Known gene expression in each type and GO enrichments are shown to the right. The graph above shows the distribution of each subclass by gestational week, and the graph below shows the subclusters of progenitors. Clusters 1–8: 204, 159, 483, 730, 397, 300, 139 and 74 cells. **b**, Dot plot for known markers of subtypes of progenitors in Fig. 2b. The size of each dot represents the percentage of cells in each cluster. Grey-to-blue gradient shows low-to-high gene expression. Progenitors: 204, 159, 483, 730, 397, 300, 139 and 74 cells. **c**, Dot plot for novel markers of subtypes of progenitors. The size of each dot represents the percentage of cells in each cluster. Grey-to-red gradient shows low-to-high gene expression. Progenitors:

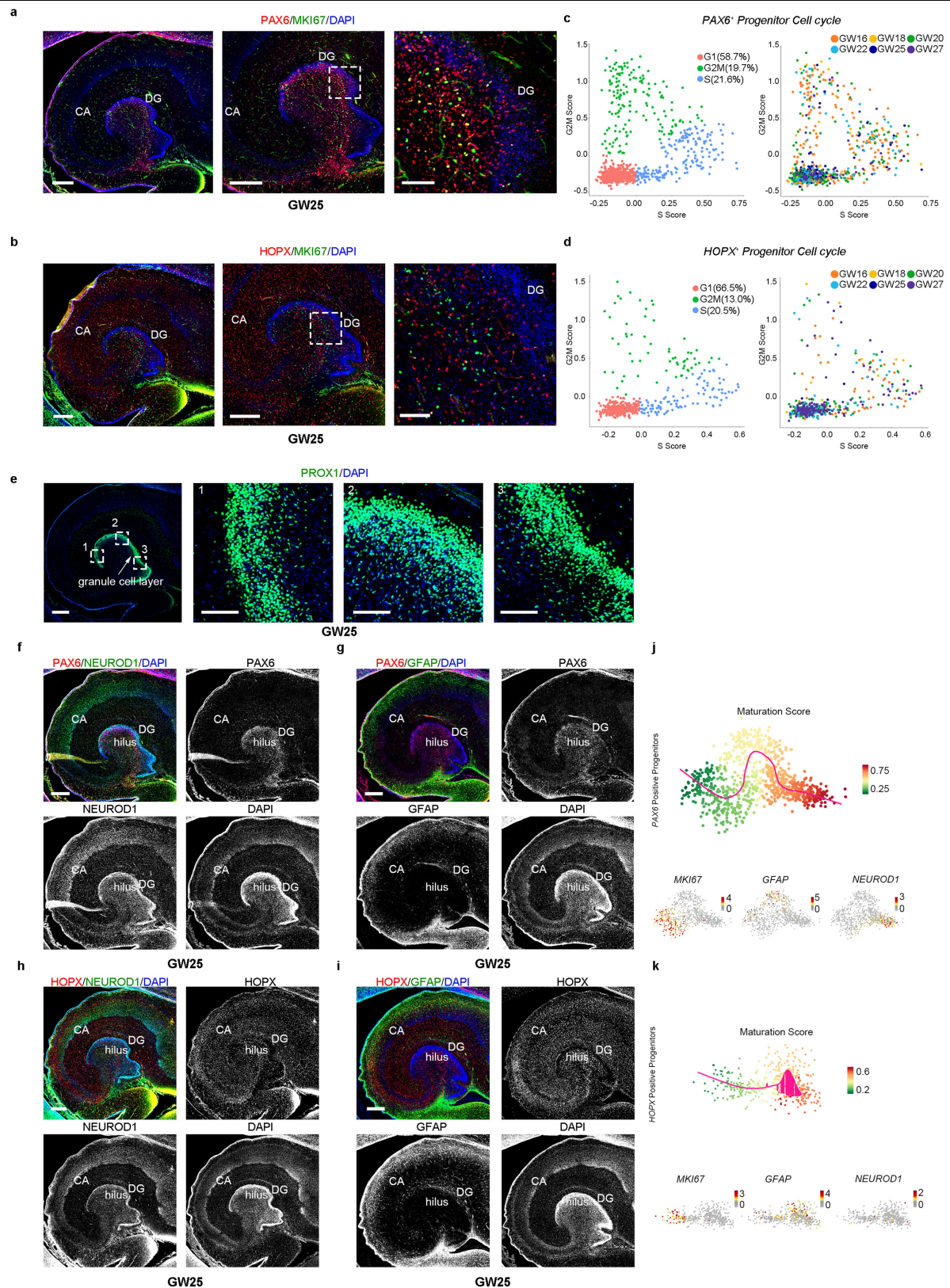
204, 159, 483, 730, 397, 300, 139 and 74 cells. **d**, Abstracted graph shows the connection on the transcriptome of different subtypes in the developing human hippocampus. Each dot represents a single cell, and cell colour represents the cell type. **e**, Abstracted graph shows the connection on the transcriptome of all subtypes in the developing human hippocampus. Each dot represents a single cell, and cell colour represents the cell type. **f**, Abstracted graph shows the connection on the transcriptome of different weeks in the developing human hippocampus. Each dot represents a single cell, and cell colour represents the week. **g, h**, Visualization of eight subtypes of progenitors in the developing human hippocampus using *t*-SNE (**g**), and expression of known markers using the same layout (**h**). Grey, no expression; red, relative expression.



**Extended Data Fig. 4 | Immunostaining of progenitors in the developing hippocampus.** **a**, Immunofluorescence images of PAX6 and SOX2 at GW11. Scale bar, 2,000  $\mu\text{m}$ . **b**, Immunofluorescence images of HOPX and SOX2 at GW11. Scale bar, 2,000  $\mu\text{m}$ . **c**, Immunofluorescence images of PROX1, PAX6,

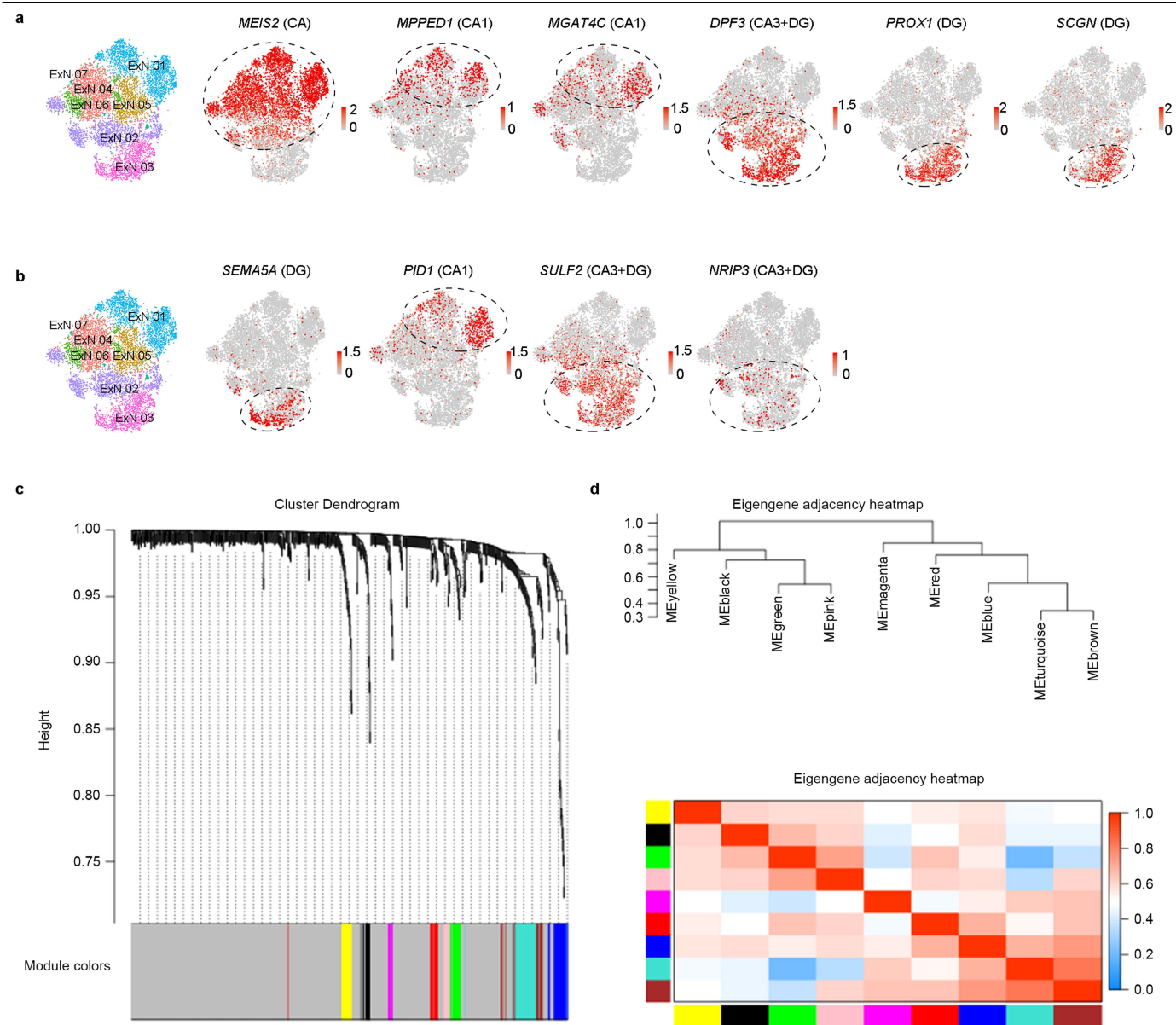
HOPX and SOX2 at GW14. Scale bar, 1,000  $\mu\text{m}$ . **d–i**, Immunofluorescence images of PROX1, PAX6, HOPX and SOX2 in GW16 (**d–f**) and GW22 (**g–i**). Scale bar, 500  $\mu\text{m}$ . The experiment was repeated three times independently with similar results.





**Extended Data Fig. 5 | Immunostaining of developing hippocampus.**  
**a, b**, Immunofluorescence images of PAX6, HOPX and MKI67 at GW25. Scale bar, 500  $\mu$ m. **c, d**, Cell cycle analysis of PAX6<sup>+</sup> (**c**) or HOPX<sup>+</sup> (**d**) progenitors. **e**, Immunofluorescence images of PROX1 in GW25 to show granule cell layer.

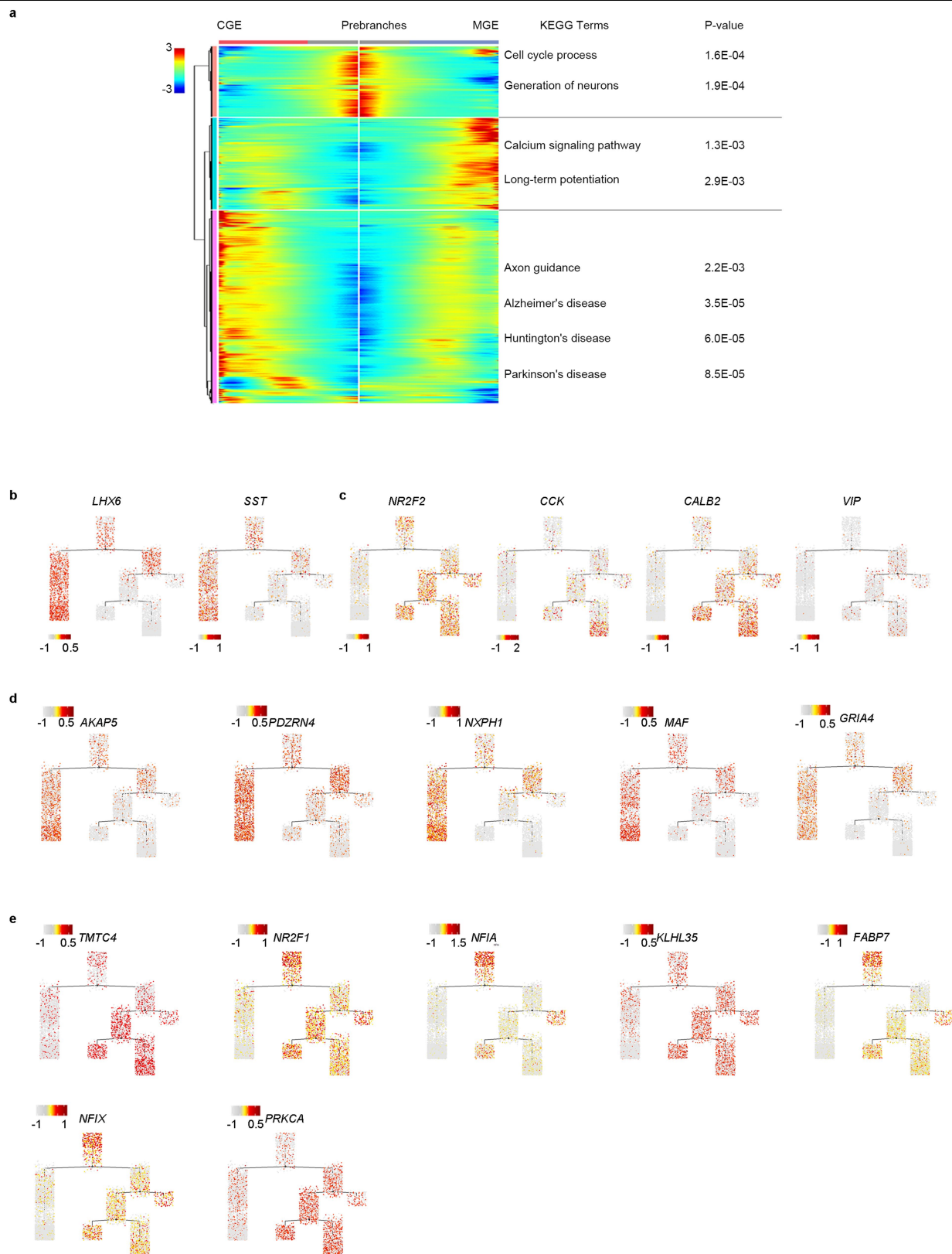
Scale bars, 500  $\mu$ m (left); 100  $\mu$ m (right, panels 1–3). **f–i**, Immunofluorescence images of PAX6, HOPX, NEUROD1 and GFAP at GW25. Scale bar, 500  $\mu$ m. The experiment was repeated three times independently with similar results. **j, k**, The maturation scores of PAX6<sup>+</sup> (**j**) and HOPX<sup>+</sup> (**k**) progenitors.



**Extended Data Fig. 6 | Molecular diversity of excitatory neurons in the hippocampus. a, b,** Expression of known markers (a) and new markers (b) shown using the same layout as in Fig. 3a. Grey, no expression; red, relative

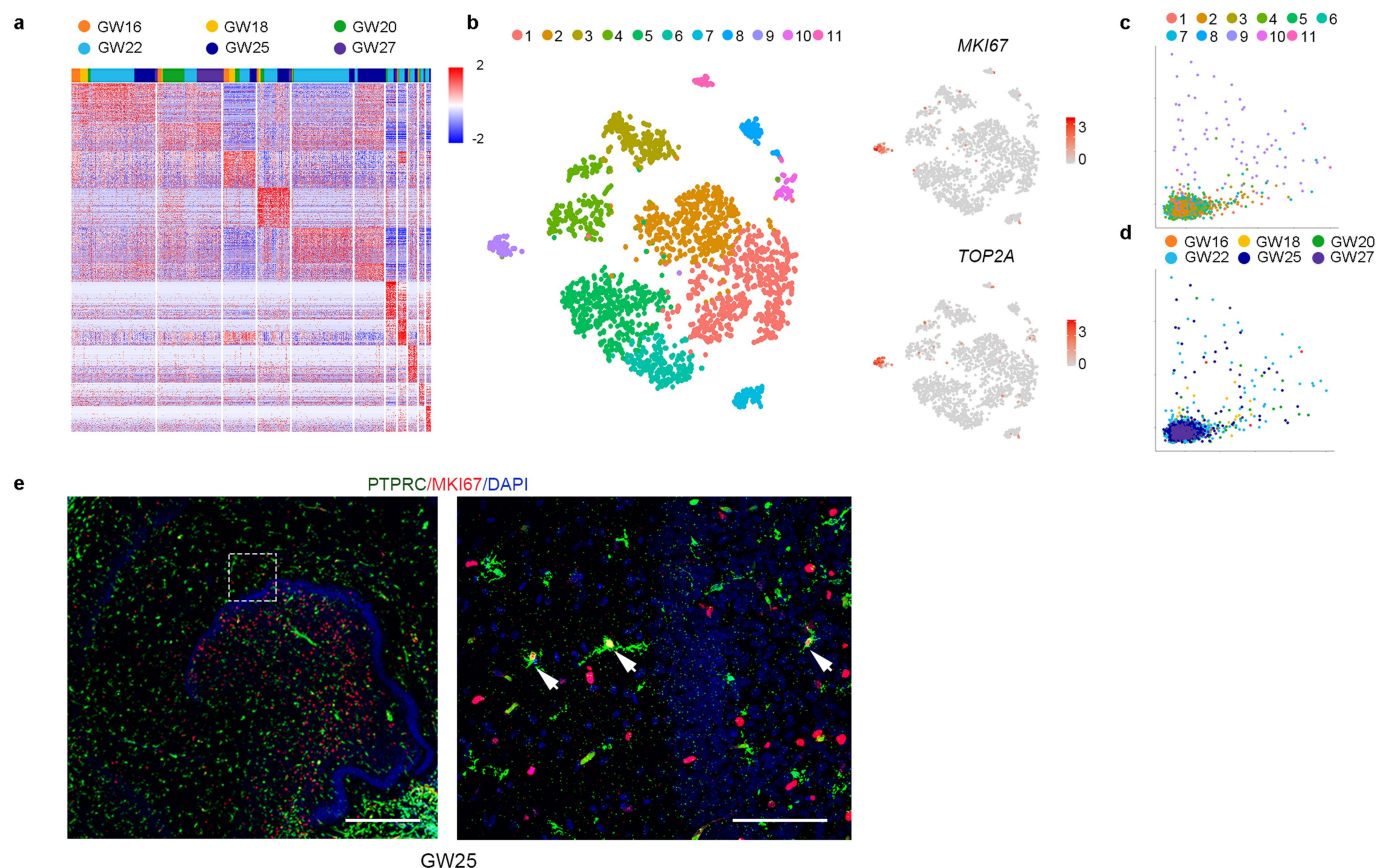
expression. **c,** Cluster dendrogram showing the modules selected to calculate the gene network in Fig. 3h. **d,** The cluster trees and heat map show the correlation of different gene modules in excitatory neurons.





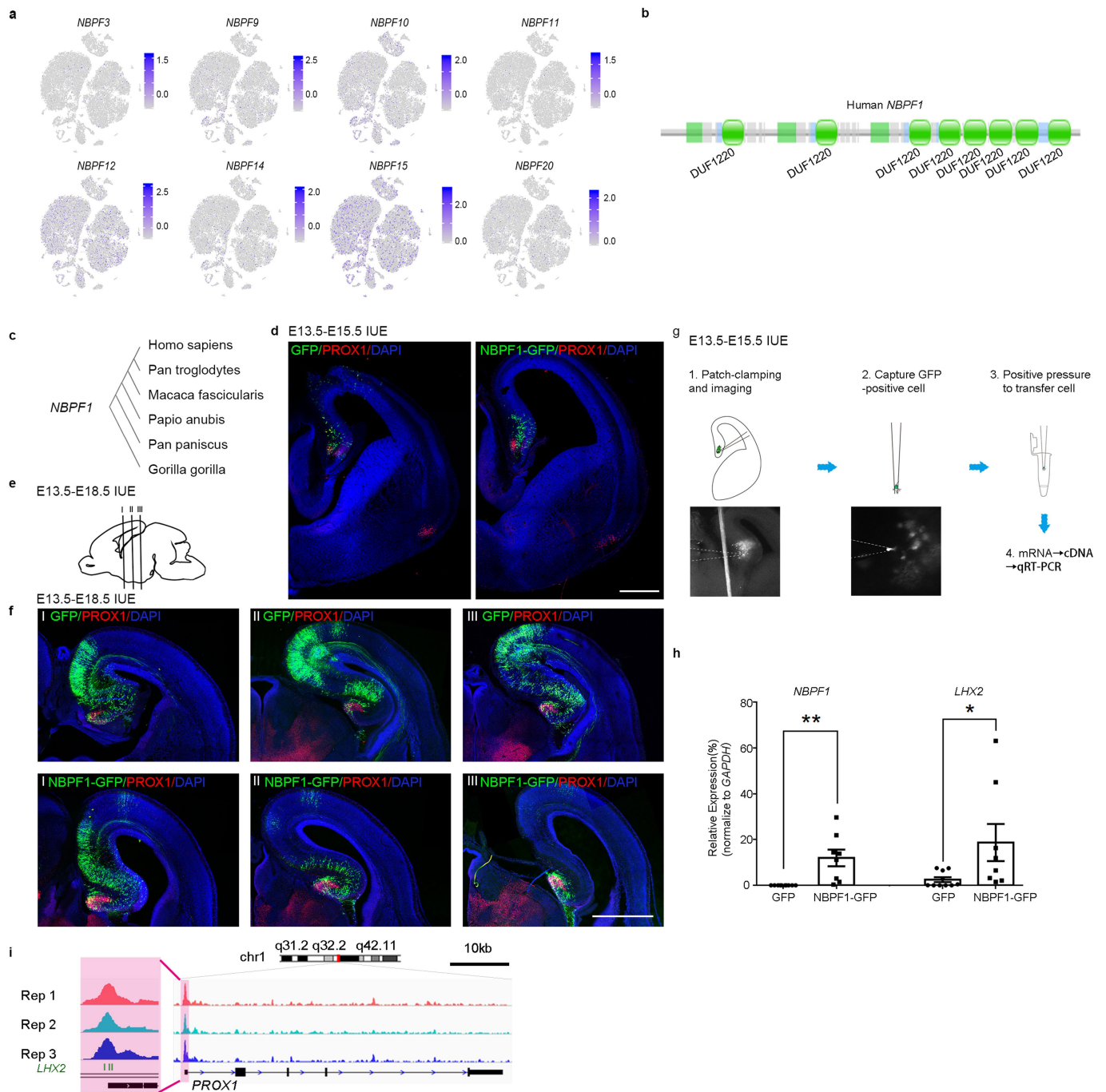
**Extended Data Fig. 7 | Molecular diversity of inhibitory neurons in the hippocampus. a,** GO analysis of modules created by clustering the two main branches from the lineage tree. The analysis reflects cell fate commitment. In this heat map, the middle represents the start of pseudo-time. From this point, one lineage moves to the CA and the other moves to the DG. Rows are GO terms correlated into different modules. Sample size: 12,115 cells. **b, c,** Expression of

known markers shown using the same layout as in Fig. 3k. **d,** Expression of novel markers of MGE-derived inhibitory neurons shown using the same layout as in Fig. 3k. Grey, no expression; red, relative expression. **e,** Expression of novel markers of CGE-derived inhibitory neurons shown using the same layout as in Fig. 3k.



**Extended Data Fig. 8 | Molecular diversity of microglia in the human hippocampus.** **a**, Heat map showing the expression levels and identities of genes in the microglia subclasses. The graph above shows the distribution of each subclass by gestational week. **b**, Visualization of ten subtypes of microglia in the developing human hippocampus using *t*-SNE. Each dot represents a single cell, and cells are laid out to show similarities. Each cell colour represents the cell type. Expression of known markers is shown using the same layout on

the right; grey, no expression; red, relative expression. Microglia: 638, 489, 246, 259, 465, 229, 84, 84, 68, 54 and 44 cells. **c**, **d**, Distribution of G1, S, and G2/M stages of the cell cycle for microglia of different subtypes (**c**) and at different gestational weeks (**d**). **e**, Immunostaining images of PTPRC and MKI67 at GW25. Scale bars, 500  $\mu$ m (left), 100  $\mu$ m (right). The experiment was repeated three times independently with similar results.



#### Extended Data Fig. 9 | NBPF family genes in the human hippocampus.

**a**, Expression of NBPF family genes shown using the same layout as in Fig. 1b. Grey, no expression; blue, relative expression. **b**, Domains of NBPF1. **c**, Evolutionary history inferred using the neighbour-joining method. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method and are in the units of the number of amino acid substitutions per site. The analysis involved six amino acid sequences. Evolutionary analyses were conducted in MEGA X. **d**, Overexpression of NBPF1 promotes DG formation at E13.5, and is observed at E15.5 in mouse. Scale bars, 500  $\mu$ m. The experiment was repeated six times

independently with similar results. **e**, Scheme depicting the position in the mouse brain at E18.5 of the slice in Fig. 1f. **f**, Overexpression of NBPF1 promotes DG formation at E13.5, and this is observed at E18.5 in mouse. Scale bars, 1,000  $\mu$ m. The experiment was repeated six times independently with similar results. **g**, Flow chart of patch-qRT-PCR. **h**, Relative expression of specific genes of GFP<sup>+</sup> cells. \*\* $P=0.0020$ , \* $P=0.0408$ , two-sided  $t$ -test;  $n=10$  GFP cells; 8 NBPF1-GFP cells. Mean  $\pm$  s.e.m. **i**, Normalized ATAC-seq profiles of *PROX1* in GW25 hippocampus with three independent biological replicates (Rep1, Rep2 and Rep3) showing the activation of *PROX1*. The amplifying panel shows the predicted LHX2 binding sites.



## Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- |                                     |                                     |  |
|-------------------------------------|-------------------------------------|--|
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | The statistical test(s) used AND whether they are one- or two-sided<br><i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/>            | A description of all covariates tested   |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | For null hypothesis testing, the test statistic (e.g. $F$ , $t$ , $r$ ) with confidence intervals, effect sizes, degrees of freedom and $P$ value noted<br><i>Give <math>P</math> values as exact values whenever suitable.</i>                            |
| <input checked="" type="checkbox"/> | <input type="checkbox"/>            | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/>            | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes   |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | Estimates of effect sizes (e.g. Cohen's $d$ , Pearson's $r$ ), indicating how they were calculated   |

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

Data collection

No software was used for data collection.

Data analysis

Custom written scripts in Cellranger v2.0.1, Python 3.7, Python package: Scanpy v1.3.8, R, R package: Monocle v2.99.0, Seurat v2.3.4, rgl v0.99.16, David 6.8, GraphPad Prism 6.0. See Methods for details on how each software is used.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

### Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The scRNA-seq data and ATAC-seq data used in this study have been deposited in the Gene Expression Omnibus (GEO) under accession numbers GSE131258.

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- ☒ Life sciences      ☐ Behavioural & social sciences      ☐ Ecological, evolutionary & environmental sciences

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	The sample size of scRNA-seq and ATAC-seq were determined by availability of human tissues. We collected 7 hippocampus from embryonic stages for scRNA-seq with 2 samples at the same developmental stage as replications . Final dataset scale was determined according to the quality control criteria as described in the methods. 6 mice were used per time point to quantify the phenotype.
Data exclusions	Cells detected with less than 800 genes were removed as low quality cells. Genes which only expressed in fewer than 30 cells (0.1% of total cell number) were excluded as recommended by Seurat (Ver.2.3.4 ).
Replication	As scRNA-seq, 2 biological replicates in GW22 and no replicates for other time points. 3 replicates were used in ATAT-seq experiments. 6 mice were used per time point to quantify the phenotype as described in figure legends. All replications were consistent for data results.
Randomization	The samples were allocated into each experimental groups based on the gestational stage. See methods 'Tissue sample collection and dissection'.
Blinding	The investigators were blinded to group allocation during data collection and analysis.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems		Methods	
n/a	Involved in the study	n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies	<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines	<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology	<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging
<input type="checkbox"/>	<input checked="" type="checkbox"/> Animals and other organisms		
<input type="checkbox"/>	<input checked="" type="checkbox"/> Human research participants		
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data		

### Antibodies

Antibodies used	For immunostaining, the following antibodies to the following proteins were used: Mouse monoclonal [MEM-28] to CD45 ,Abcam,ab8216, GR302332-1; Rabbit polyclonal to NEUROD2,Abcam,ab104430,GR94291-4; Rabbit Polyclonal to PAX6,BioLegend,901301,B201255; Goat polyclonal to SOX2,Santa Cruz,sc-17320,H1406; Mouse monoclonal to NEUROD1, Abcam,ab60704,GR3183945-2; Rabbit monoclonal [EPR18114] to HMGA2, Abcam, ab207301; Rabbit polyclonal to HOPX, Santa Cruz, sc-30216, D1615; Mouse monoclonal [B56] to Ki67, BD Biosciences, 550609, 19679; Rabbit monoclonal [EPR19273] to PROX1, Abcam, ab199359, GR45436-1; Rabbit polyclonal to OLIG2, Millipore, ab9610,3018858 ; Human monoclonal [IGX3421] to Myelin Basic Protein, Abcam, ab209328,GR278417-4; Mouse monoclonal [GA5] to GFAP,CST,3670S,6.
Validation	All antibodies were validated by the supplier for human or mouse samples and by comparing to the manufacturer's or in-house results.

### Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals	CD1 mouse, female, timed pregnant at E13.5.
Wild animals	This study did not involved the wild animals.
Field-collected samples	This study did not involved the samples collected from the field.

## Ethics oversight

Mouse housing and experimental protocols in this study were in compliance with the guidelines of the Institutional Animal Care and Use Committee of the Institute of Biophysics, CAS.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Human research participants

Policy information about [studies involving human research participants](#)

## Population characteristics

All embryo and fetal tissues were between 16-27 gestational weeks. Gestational age was measured in weeks from the first day of the woman's last menstrual cycle to the sample collecting date.

## Recruitment

Beijing Anzhen Hospital was in charge of recruiting donors for this research. The patients decided to have an abortion first, and then they were asked whether they would agree to donate the fetal tissues to this study. The de-identified human fetal tissue samples were collected after the donor patients signing informed consent document. The tissue collection and research protocols were approved by the Reproductive Study Ethics Committee of Beijing Anzhen Hospital and the institutional review board (ethics committee) of the Institute of Biophysics.

## Ethics oversight

The Reproductive Study Ethics Committee of Beijing Anzhen Hospital ; the institutional review board (ethics committee) of the Institute of Biophysics.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

# A developmental landscape of 3D-cultured human pre-gastrulation embryos

<https://doi.org/10.1038/s41586-019-1875-y>

Received: 24 December 2018

Accepted: 5 December 2019

Published online: 12 December 2019

Lifeng Xiang<sup>1,2,3,4,7</sup>, Yu Yin<sup>1,3,4,7</sup>, Yun Zheng<sup>1,4,5,7</sup>, Yanping Ma<sup>2,7</sup>, Yonggang Li<sup>2,7</sup>, Zhigang Zhao<sup>1</sup>, Junqiang Guo<sup>1,5</sup>, Zongyong Ai<sup>1,4</sup>, Yuyu Niu<sup>1,4</sup>, Kui Duan<sup>1,4</sup>, Jingjing He<sup>1,4</sup>, Shuchao Ren<sup>1</sup>, Dan Wu<sup>1</sup>, Yun Bai<sup>2</sup>, Zhouchun Shang<sup>6</sup>, Xi Dai<sup>6</sup>, Weizhi Ji<sup>1,4\*</sup> & Tianqing Li<sup>1,4\*</sup>

Our understanding of how human embryos develop before gastrulation, including spatial self-organization and cell type ontogeny, remains limited by available two-dimensional technological platforms<sup>1,2</sup> that do not recapitulate the *in vivo* conditions<sup>3–5</sup>. Here we report a three-dimensional (3D) blastocyst-culture system that enables human blastocyst development up to the primitive streak anlage stage. These 3D embryos mimic developmental landmarks and 3D architectures *in vivo*, including the embryonic disc, amnion, basement membrane, primary and primate unique secondary yolk sac, formation of anterior–posterior polarity and primitive streak anlage. Using single-cell transcriptome profiling, we delineate ontology and regulatory networks that underlie the segregation of epiblast, primitive endoderm and trophoblast. Compared with epiblasts, the amniotic epithelium shows unique and characteristic phenotypes. After implantation, specific pathways and transcription factors trigger the differentiation of cytotrophoblasts, extravillous cytotrophoblasts and syncytiotrophoblasts. Epiblasts undergo a transition to pluripotency upon implantation, and the transcriptome of these cells is maintained until the generation of the primitive streak anlage. These developmental processes are driven by different pluripotency factors. Together, findings from our 3D-culture approach help to determine the molecular and morphogenetic developmental landscape that occurs during human embryogenesis.

Technical limitations preclude the precise delineation of early human embryogenesis, such as architecture formation and cell-type specification. Recent *in vitro* implantation platforms using a two-dimensional (2D) culture approach have revealed some developmental landmarks of early human embryos *in vivo*<sup>1,2</sup>. However, these 2D culture embryos are largely flattened, which creates an imperfect model of normal 3D embryonic development *in vivo* and limits classification using equivalent Carnegie stages<sup>3–5</sup>. Although pluripotent stem cells can model some phenotypes of the human amnion sac, amniogenesis or organizer<sup>6–8</sup>, we still cannot authentically mimic human embryonic development, especially for embryo lineage ontogeny. These existing methods exclude crosstalk among different cell types in embryos. Here, we report a 3D-culture system that enables development of the human blastocyst up to the primitive streak anlage (PSA) stage. Using the 3D platform, we reveal a developmental landscape of human pre-gastrulation embryos.

## 3D architectures in 3D-cultured embryos

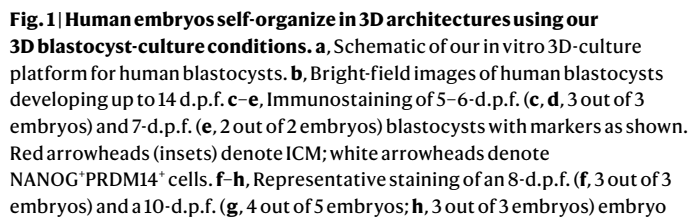
We used donated human embryos from surplus embryos after clinical *in vitro* fertilization. We tested whether the culture media used for 2D human embryo cultures—IVC1 and IVC2<sup>1,2</sup>—were suitable for 3D

culture of blastocysts. These media sustained only 6.3% of embryos until 14 days post-fertilization (d.p.f.) using morphological embryo observations (Extended Data Fig. 1a–f). We then added sodium lactate, sodium pyruvate and ROCK inhibitor (Y27632)<sup>9</sup> to IVC1 and IVC2 media, resulting in modified IVC1 (mIVC1) and mIVC2, respectively. Cultures in mIVC1 and mIVC2 sustained 23.4% of human blastocysts until 14 d.p.f. (Extended Data Fig. 1a–f).

We then designed a series of 3D extracellular matrix with Matrigel embedding to identify an ideal 3D blastocyst-culture system. We verified embryonic development using morphological observations and staining for specific lineage markers—OCT4 for the inner cell mass (ICM) and epiblast (EPI), GATA6 for primitive endoderm/hypoblast (PrE) and CK7 for trophoblast (TrB) (Extended Data Fig. 1g–j). We found that 10% Matrigel yielded the best outcome and enabled 23.5% of blastocysts to develop to 14 d.p.f. with normal embryonic structures (Extended Data Fig. 1g–m). We used 10% Matrigel in conjunction with mIVC1 and mIVC2 to culture blastocysts unless noted otherwise (Fig. 1a).

Morphological observations revealed that we could culture human blastocysts up to 14 d.p.f. using our 3D platform (Fig. 1b). Nearly all blastocysts at 5–6 d.p.f. were positive for GATA6, but negative for CK7, whereas the ICM expressed OCT4, NANOG, KLF17 and PRDM14 (Fig. 1c, d, Extended Data Fig. 9b). At 7 or 8 d.p.f., GATA6, CK7 and OCT4 showed

<sup>1</sup>Yunnan Key Laboratory of Primate Biomedical Research, Institute of Primate Translational Medicine, Kunming University of Science and Technology, Kunming, China. <sup>2</sup>Department of Reproductive Medicine, The First People's Hospital of Yunnan Province, Kunming, China. <sup>3</sup>Faculty of Environmental Science and Engineering, Kunming University of Science and Technology, Kunming, China. <sup>4</sup>Yunnan Provincial Academy of Science and Technology, Kunming, China. <sup>5</sup>Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming, China. <sup>6</sup>BGI-Shenzhen, Shenzhen, China. <sup>7</sup>These authors contributed equally: Lifeng Xiang, Yu Yin, Yun Zheng, Yanping Ma, Yonggang Li. \*e-mail: wji@lpr.cn; litq@lpr.cn

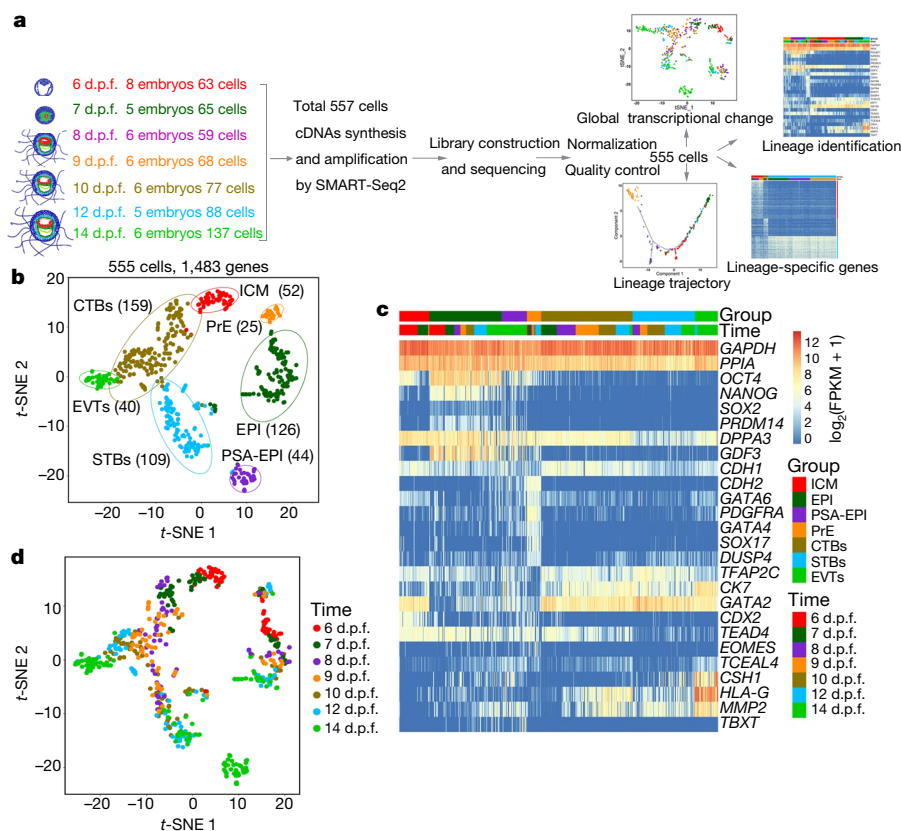


(Extended Data Fig. 2). The white and red arrows in **f** denote epiblast and primary yolk sac, respectively. The white arrowhead in **g** denotes the amniotic cavity. **i**, Representative staining of a 12-d.p.f. embryo (6 out of 8 embryos). Red arrowheads denote EPIs; white arrows denote AME. **j**, Representative staining of 14-d.p.f. embryos (4 out of 6 embryos). **k**, Magnification of the square shown in **j**, rotated 90° clockwise. AC, amniotic cavity; PE, parietal endoderm (white arrows) (k2); VE, visceral endoderm (red arrows) (k1). Scale bars, 100  $\mu$ m (**b**), 50  $\mu$ m (**c–g**, **i**, **j**) or 10  $\mu$ m (**h**). See also Extended Data Figs. 1, 2.

(Fig. 1k), which suggests that it arises from PrEs that rapidly give rise to the visceral endoderm and parietal endoderm after implantation<sup>14</sup>. Embryos displayed a 3D spherical structure with a disc-shaped bilaminar structure, SYS and amnion, and with the TrBs on the amnion side of the embryo initiating trophoblast differentiation and undergoing spatial and asymmetric development to concentrate at the amnion side (Extended Data Fig. 1n–q, Supplementary Videos 1, 2). Together, our 3D platform can support self-organization with spatial architectures of human embryos.

We performed 557 single-cell RNA sequencing (scRNA-seq) from 42 embryos at 7 developmental stages to scan the transcriptome of our 3D cultured embryos (Fig. 2a). Following quality control and stringent filtering, we used 555 single cells with 23,270 genes for subsequent analyses (Extended Data Fig. 3a–c). *t*-distributed stochastic neighbour embedding (*t*-SNE) analyses revealed seven clusters, identified as ICM, EPI, PrE, TrB (including cytotrophoblasts (CTBs), syncytiotrophoblasts (STBs) and extravillous cytotrophoblasts (EVTs)) and PSA-EPI based on lineage-specific marker expression and developmental time (Fig. 2b–d). Continuous transcriptome shifts from 6 to 14 d.p.f. reflected a transition from pre- to post-implantation (Fig. 2d). Integrated analysis of scRNA-seq data from different embryo sources<sup>15–18</sup> revealed that





**Fig. 2 | Lineage delineation by transcriptome using scRNA-seq. a**, Schematic of single-cell collection and transcriptome analyses. **b–d**, *t*-SNE analyses revealed seven clusters, identified as the ICM, EPI, PrE, TrB (including CTBs,

STBs and EVTs) and PSA-EPI based on classical lineage-specific marker expression (c) and developmental time (d). FPKM, fragments per kilobase of transcript per million mapped reads. See also Extended Data Fig. 3.

lineage segregation occurred—but incompletely—in human 6-d.p.f. blastocysts independent of sample sources across all samples, and cell fates appeared fixated in 7–9-d.p.f. embryos (Extended Data Fig. 3d, e). Thus, we used scRNA-seq in 7–9-d.p.f. 3D embryos to examine the regulators that segregate TrBs, PrEs and EPIs. We observed EPI-, PrE- and TrB-specific genes associated with different signalling pathways and transcription factors (Extended Data Fig. 3f–j, Supplementary Table 1). Lineage-specific gene comparison between previous results<sup>18</sup> and our results showed that core lineage transcription factors are maintained across different samples, while some differences in gene expression may be contributed by different developmental stages of embryos (Extended Data Fig. 3k, Supplementary Table 1.6).

## AME–EPI separation

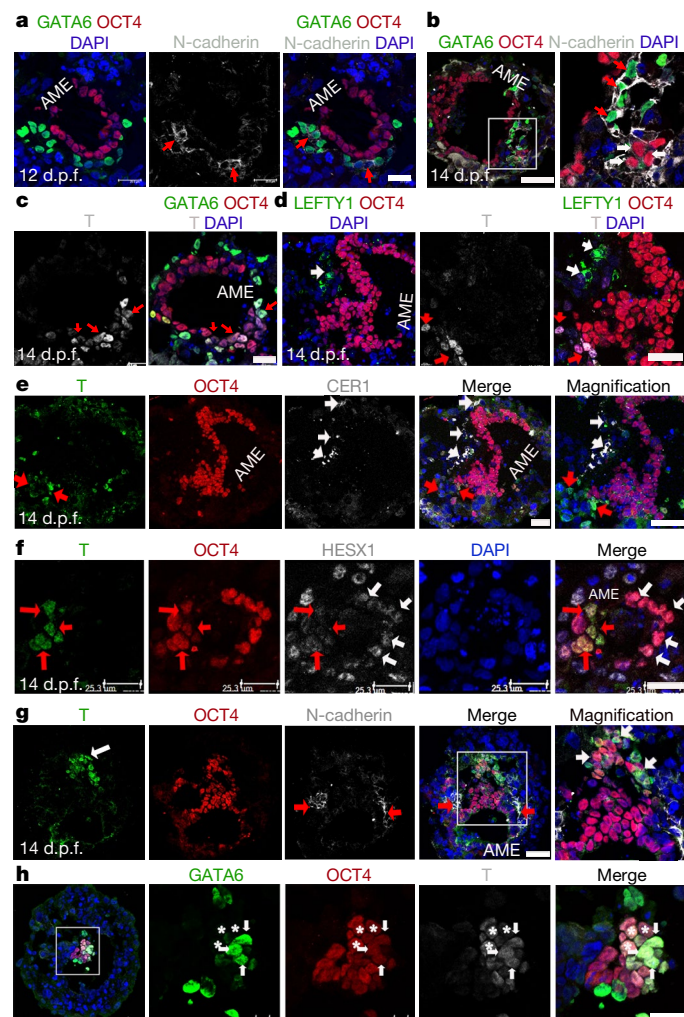
In contrast to mouse EPI, the proximal EPI in primates segregates an additional AME lineage before gastrulation<sup>19</sup>. However, human AME remains unclear in the absence of good molecular markers. As tissue morphogenesis requires cell–cell adhesion proteins<sup>20</sup>, we observed a symmetrical distribution of E-cadherin (also known as CDH1) on the cell membrane of columnar EPIs, while squamous AME displayed no to weak E-cadherin expression (Extended Data Fig. 4a, b). However, widespread E-cadherin distributions in human EPIs were concentrated on the apical site of wedge-shaped mouse EPIs<sup>21</sup>. Amnion formation requires signalling from the basement membrane generated by visceral endoderm<sup>21,22</sup>. The layer between PrEs and EPIs formed a laminin-containing basement membrane, enveloping EPIs but not AME (Extended Data Fig. 4c, d). The AME weakly expressed OCT4, NANOG and SOX2 (Extended Data Fig. 4b, f). At 6 and 8 d.p.f., E-cadherin was ubiquitously distributed on EPI and TrB cell membranes, whereas laminin enveloped the entire EPI cluster and was widely expressed in TrBs (Extended Data Fig. 4g–j).

At 10 d.p.f., laminin concentrated around EPIs to form the basement membrane but was lost in the AME (Extended Data Fig. 4k, l).

We next checked expressions of ERIZN and WGA, which localize to the apical surfaces of human-pluripotent-stem-cell-derived amnion<sup>6,8</sup>. ERIZN equally contributed to the apical surfaces of EPIs and AME (Extended Data Fig. 4m). However, WGA expressed in extra-embryonic cells, but not in EPIs and AME (Extended Data Fig. 4n), which indicates differences between 3D-embryo- and human-pluripotent-stem-cell-derived amnion. Given that obvious separation of AME–EPI occurred at 12 and 14 d.p.f., we analysed scRNA-seq from 12- and 14-d.p.f. EPIs in the post- and PSA-EPI clusters (Fig. 2b). The *t*-SNE analysis classified these EPIs into three clusters, termed AME, intermediate state cells and EPIs, on the basis of their gene-expression profiles (Extended Data Fig. 4o, p). Compared to EPIs, AME significantly downregulated pluripotency genes and upregulated genes expressed in the AME of 12–17-d.p.f. monkey embryos (*TFAP2C*, *MSX2* and *BMP4*)<sup>23</sup> or self-organized amnion from human pluripotent stem cells (*TFAP2A* and *GATA3*)<sup>6,8</sup> (Extended Data Fig. 4p). High expression of hormone genes in AME (Extended Data Fig. 4q–u, Supplementary Table 2) corresponds to the AME of human placentas producing human chorionic gonadotropin (hCG)<sup>24</sup>. These results indicate the AME is a distinct population with specific phenotypes compared to EPIs.

## Forming anterior–posterior polarity and PSA

Primitive streak remains poorly defined and remains an enigmatic structure in human embryos. One hallmark of primitive streak formation is the epithelial–mesenchymal transition and upregulated N-cadherin (CDH2), a mesenchymal marker<sup>25</sup>. We found N-cadherin localized in PrEs before 12 d.p.f. and was activated in some OCT4-expressing cells outside the EPI or near the AME–EPI junction at 14 d.p.f. (Fig. 3a, b). The result was confirmed by scRNA-seq data (Fig. 2c, Extended Data Fig. 5l).



**Fig. 3 | Human embryos display anterior–posterior polarity and formation of PSA.** **a, b**, Staining of embryo sections at 12 d.p.f. (**a**, 3 out of 3 embryos) and 14 d.p.f. (**b**, 4 out of 5 embryos) with N-cadherin. Red arrows denote N-cadherin<sup>+</sup> PrEs; white arrows denote N-cadherin<sup>+</sup> EPIs migrating out from the disc (**b**). **c–f**, Staining of embryo sections at 14 d.p.f. with T, LEFTY1, CER1 and HESX1. In **c**, red arrows denote T<sup>+</sup> cells (5 out of 8 embryos). In **d**, white arrows denote LEFTY1<sup>+</sup> and red arrows denote T<sup>+</sup> cells (3 out of 4 embryos). In **e**, white arrows denote CER1<sup>+</sup> cells; red arrows denote T<sup>+</sup> cells (**e**, 2 out of 2 embryos). In **f**, white arrows denote T<sup>+</sup> HESX1<sup>+</sup> EPIs; red arrows denote T<sup>+</sup> HESX1<sup>−</sup> EPIs (2 out of 2 embryos). **g, h**, Staining of human embryo sections at 14 d.p.f. In **g**, white arrows denote T<sup>+</sup> cells; red arrows denote N-cadherin<sup>+</sup> cells (4 out of 6 embryos). In **h**, white arrows denote T<sup>+</sup> OCT4<sup>+</sup> GATA6<sup>+</sup> cells; asterisks denote T<sup>+</sup> OCT4<sup>+</sup> GATA6<sup>−</sup> cells (3 out of 4 embryos). Scale bars, 50  $\mu$ m (**a–e, g**) or 25  $\mu$ m (**f, h**). See also Extended Data Figs. 5–7.

During embryonic development in mice, the anterior visceral endoderm secretes LEFTY1 and CER1 to antagonize posteriorizing morphogens, and OTX2 regulates the anterior visceral endoderm and anterior–posterior axis<sup>22</sup>. We observed some EPIs expressing T-box transcription factor T (T, also known as brachyury or TBXT), an early marker for primitive streak<sup>26</sup>, but repressing OCT4 at the EPIs near the AME compartment boundary at 14 d.p.f. but not at 12 d.p.f. (Fig. 3c, Extended Data Fig. 5a). Although we did not observe mutually exclusive expression of SOX2 (anterior commitment) and NANOG (posterior commitment)<sup>27</sup> in EPIs, we observed CER1, LEFTY1 and OTX2 expression on one side of the 14-d.p.f. embryonic disc (Fig. 3d, e, Extended Data Fig. 5b–e), which suggests formation of the anterior visceral endoderm. In contrast to expression in mouse embryonic cells but not the extra-embryonic compartment<sup>22</sup>, OTX2 was expressed in the PrEs of our 3D

embryos (Extended Data Figs. 3f, 5e). The location of T-expressing EPIs was opposite to the anterior visceral endoderm side (Fig. 3d, e).

Next, we examined the expression of HESX1, as a marker of EPI<sup>28</sup>, early anterior differentiation or visceral endoderm<sup>29</sup>. We observed that HESX1 was uniformly expressed in 12-d.p.f. EPIs (Extended Data Fig. 5f). Some EPIs lost HESX1 but upregulated T expression at 14 d.p.f. (Fig. 3f). Furthermore, T<sup>+</sup> HESX1<sup>+</sup> and T<sup>+</sup> HESX1<sup>−</sup> EPIs localized on opposite sides of the EPIs, representing anterior and posterior EPIs. scRNA-seq revealed decreased expression of *HESX1* in the PSA-EPIs (14 d.p.f.) (Extended Data Fig. 5g). EPIs from 14-d.p.f. embryos branched to produce 2 populations: T<sup>+</sup> HESX1<sup>+</sup> and T<sup>+</sup> HESX1<sup>−</sup> cells (Extended Data Fig. 5h), which further defines anterior and posterior cell fate. T<sup>+</sup> HESX1<sup>−</sup> EPIs enriched with posterior and primitive streak genes, whereas T<sup>+</sup> HESX1<sup>+</sup> EPIs showed upregulated early anterior genes (Extended Data Fig. 5i, Supplementary Table 3).

Using serial sections of whole embryos, we observed that some 14-d.p.f. embryos formed a cell dissemination region<sup>30</sup>, in which some T<sup>+</sup> EPIs focally migrated from the embryonic disc, disrupted barriers formed by PrE (with N-cadherin expression), invaded the space near the visceral endoderm and co-expressed GATA6 (Fig. 3g, h, Extended Data Figs. 6, 7). FLK1 (also known as KDR), a marker of extra-embryonic mesoderm<sup>31</sup>, was highly expressed in the EPIs, but lost in T<sup>+</sup> cells migrating from the embryonic disc (Extended Data Fig. 5j), which suggests that the latter is not extra-embryonic mesoderm. Some FLK1<sup>+</sup> cells localized between PrEs and TrBs, indicating extra-embryonic mesenchyme (Extended Data Fig. 5k). Compared with post-implantation EPIs, PSA-EPIs significantly upregulated primitive streak genes (Extended Data Fig. 5l). However, the absence of neural gene expression (Extended Data Fig. 5m–o) suggests that the 14-d.p.f. embryos have not developed to generate the initial nervous system yet. We conclude that our 14-d.p.f. embryos were at the PSA stage, which meets the internationally recognized ethical limit for human embryo culture. Together, we conclude that 1 population of 14-d.p.f. EPIs underwent changes and has initiated anterior–posterior polarity and primitive streak formation.

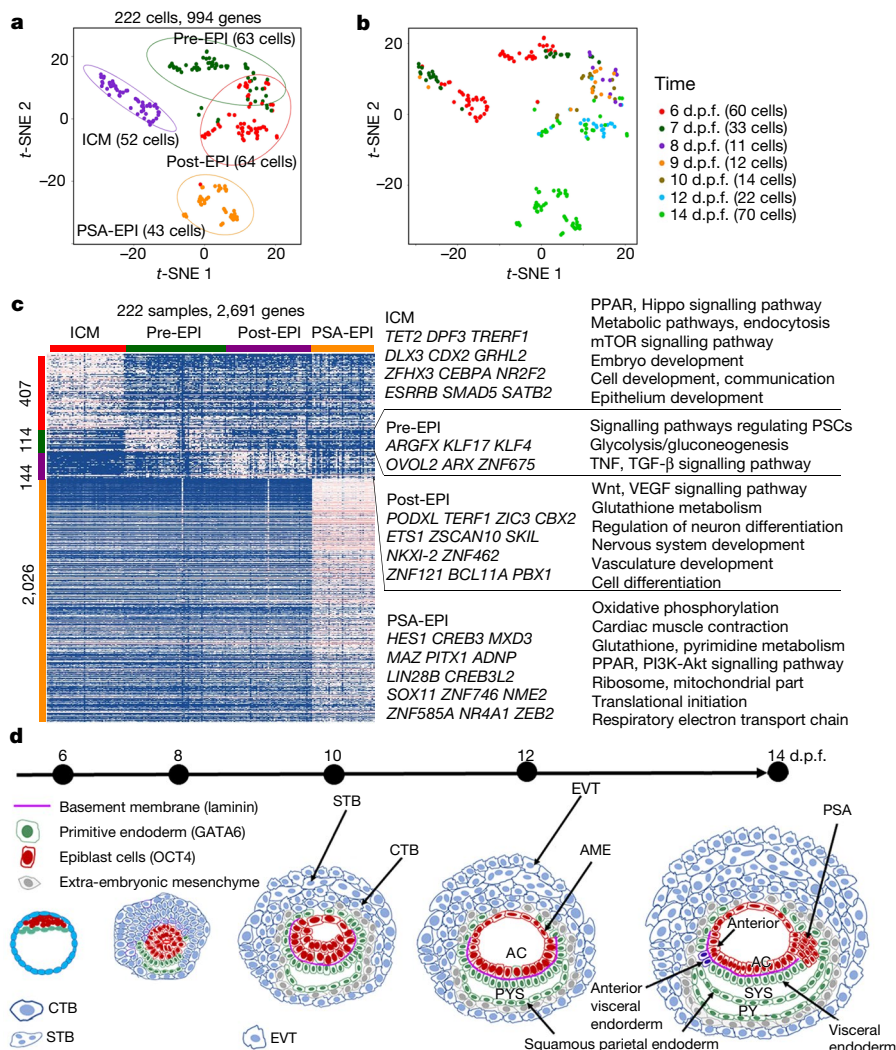
Continuous cell proliferation is key to evaluate embryonic developmental status. Cell proliferation in 2D-cultured embryos occurs only within 8–10 d.p.f., but not after 10 d.p.f.<sup>2</sup>, which implies that 2D embryos do not survive much beyond 14 days. Our 3D embryos maintained continuous proliferation of EPIs, TrBs and PrEs at all stages (Extended Data Fig. 5p–r). We predict that continuous cell proliferation may facilitate human embryos to develop beyond 14 d.p.f. to initiate gastrulation.

## Development of the trophoblast lineage

The human placenta consists of three major TrB subpopulations: CTBs, EVTs and STBs. In our 3D-cultured embryos, the TrBs surrounding EPIs and PrEs presented a polarized localization of F-actin (Extended Data Fig. 8a). CK7<sup>+</sup> cells near the EPI–PrE bilayer had a single nucleus that expressed TEAD4 and E-cadherin (Extended Data Fig. 8a–c), which indicate a CTB identity. Multinucleated cells in the layer adjacent to AME expressed hCG (Extended Data Fig. 8d, e), which suggests STB identity. One outer-layer population highly expressed HLA-G, a marker of EVTs, in 12-d.p.f. embryos, which significantly increased in 14-d.p.f. embryos (Extended Data Fig. 8f, g). Mutually exclusive expression of hCG, TEAD4 and HLA-G in most cells showed molecular and physical delineation of three TrB types at 12–14 d.p.f. (Extended Data Fig. 8h, i).

We identified the top 2,603 most variable genes across the ICM and all TrBs using scRNA-seq data. On the basis of developmental time and markers, *t*-SNE analysis categorized them into six populations: pre-CTBs (TEAD4<sup>+</sup>HLA-G<sup>−</sup>), post-CTBs, early STBs (CGB<sup>+</sup>CSH1<sup>+</sup>HLA-G<sup>low</sup>), STBs, early EVTs (HLG-A<sup>+</sup>CSH1<sup>+</sup>MMP2<sup>+</sup>ERBB2<sup>+</sup>) and EVTs<sup>32,33</sup> (Extended Data Fig. 8j–l). Continuous transcriptome shifts from 6 to 14 d.p.f. revealed a stepwise developmental progression, in which CTBs produce EVTs and STBs, and segregation initiated at 9 d.p.f. and completed at 12 d.p.f. (Extended Data Fig. 8l, m). Notably, *CDX2* expression quickly





**Fig. 4 | Development of EPI lineage.** **a, b**, t-SNE analyses revealed four clusters, identified as ICM, pre-EPI, post-EPI and PSA-EPI based on developmental time. **c**, Heat map of genes specific for every cell type. Representative transcriptional factors and KEGG pathways are shown (Supplementary Table 5). PSC, pluripotent stem cells. **d**, Model of human pre-gastrulation embryo development landmarks based on our results and the Carnegie series. See also Extended Data Fig. 9.

decreased after TrB specification at 7 d.p.f. (Extended Data Fig. 8k), given the absence of *CDX2* expression in human peri-implantation embryos and trophoblast stem cells<sup>33</sup>.

Because the human placenta must secrete steroids or polypeptide hormones to maintain pregnancy, we analysed the expression of 120 polypeptide hormone genes produced by TrBs<sup>32</sup> (Extended Data Fig. 8n). STBs, which primarily produce placental hormones, expressed 31 hormone genes starting from 8 d.p.f. EVTs expressed 19 polypeptide hormone genes with upregulated expression over culture, which indicates maturation. Expression of *CGA*, *PGF* and *CGB* family genes in EVTs suggested pre-gastrulation EVTs can secrete hCG, progesterone and oestrogen. However, these genes are significantly downregulated in 8-week EVTs, and completely disappear in 24-week EVTs<sup>32</sup>. Thus, the ability of EVTs to secrete hormones gradually decreases over placenta development.

Next, we identified genes corresponding to the different TrB subtypes (Extended Data Fig. 8o). CTB-, STB- and EVT-specific genes closely associated with their functions and specific signalling pathways accordingly to their characteristics<sup>33</sup> (Supplementary Table 4). EVT-specific genes helped to regulate the immune system and angiogenesis. These results corroborate the finding that EVTs in human first-trimester placentas are crucial for immunomodulatory and spiral artery remodelling of the early maternal–fetal interface<sup>34</sup>. We determined the top-ranked transcription factors that control TrB development. Transcription factors for CTBs, STBs and EVTs included well-documented TrB and pluripotency factors and new potential transcription factors, such as *MYBL2*, *TCF7L1* and *NR2F2* (Extended Data Fig. 8o).

## Epiblast development and transition

We analysed EPI transcriptome dynamics across development, which revealed four main clusters: ICM, pre-implantation EPI (pre-EPI), post-EPI and PSA-EPI (Fig. 4a, b). When tracking naive and primed pluripotent gene expression, we found that embryos lost naive genes *TFCP2L1*, *KLF17* and *KLF4* and activated primed gene *CD24* after implantation, while maintaining general pluripotent genes (Extended Data Fig. 9a–d). scRNA-seq data confirmed EPI pluripotent state transition (EPST)<sup>35,36</sup> (Extended Data Fig. 9e, f).

We speculated whether EPIs from different developmental stages show distinct pluripotency regulatory networks by performing PluriNetWork analysis using existing mouse databases<sup>37,38</sup>. Different combinations of key pluripotency regulators dominated and coordinated EPST networks (Extended Data Fig. 9g–j). Naive pluripotency genes (*ESRRB*, *KLF4* and *TFCP2L1*) only occupied ICM networks, which suggests that human EPIs quickly lose naive pluripotency after lineage diversification, consistent with monkey EPIs that only transiently express naive pluripotency before the late- and hatched-blastocyst stages<sup>38</sup>. Genes specific for different developmental stage EPIs revealed that EPIs maintained a stable transcriptome from pre-implantation to post-implantation with marked changes in gene expression occurring during the pre-EPI transition and PSA initiation stage (Fig. 4c). Different EPI pluripotency states were dominated by different transcription factors and regulatory pathways (Fig. 4c, Supplementary Table 5). Pairwise comparisons showed similar data (Extended Data Fig. 9k–n, Supplementary Table 6).

To compare cynomolgus and human EPI development, we analysed scRNA-seq from monkey 6–17-d.p.f. EPIs<sup>39</sup>. Principal component analysis (PCA) revealed that cynomolgus and human cells separated along the PC1, representing a major species difference (Extended Data Fig. 10a). We identified genes with significantly positive or negative PC1 scores for cynomolgus and human genes (Extended Data Fig. 10b, Supplementary Table 7). Along the PC2 and PC3 axes, cynomolgus and human EPIs cells were plotted to reflect their similar developmental transitions and conserved signalling pathways and Gene Ontology (GO) terms (Extended Data Fig. 10a, c, Supplementary Table 7). To examine species differences, we compared the dynamics of naive genes and signalling pathways over EPST in human and reported monkey data<sup>39</sup>. In contrast to human EPST, *TBX3* and *SOX15* gradually decreased, and *UTF1* and *NROB1* were absent during monkey EPST<sup>39</sup> (Fig. 4c, Extended Data Fig. 10d, e). Pathway analysis for upregulated genes during the EPST revealed a similar enrichment trend for NOTCH, BMP and FGF signalling pathways in humans and monkeys<sup>39</sup> (Extended Data Fig. 10f–h). However, *LEFTY1*, *LEFTY2* and *NODAL* in BMP signalling displayed different patterns between human and monkey EPST (Extended Data Fig. 10f). Together, human and monkey EPIs had unique phenotypes and similarities during development.

## Discussion

Here, we report a 3D-culture system that successfully cultured human blastocyst growing to the PSA stage. These 3D embryos can recapitulate almost all key 3D architectures and developmental landmarks of in vivo pre-gastrulation embryos. By contrast, many 3D structures and developmental landmarks—such as AME–EPI separation, basement membrane, SYS, anterior visceral endoderm, anterior–posterior polarity initiation and PSA—were not found using in vitro implantation platforms of 2D-cultured embryos<sup>1,2</sup> and using human pluripotent stem cells that model early developmental embryonic events<sup>6–8</sup>. Our 3D-cultured embryos recapitulated the timing and outcome of lineage segregation and development, which more authentically mimicked early human embryonic development in vivo (Fig. 4d).

Because human embryogenesis is not well understood, we established a platform to delineate EPIs, AMEs, PrEs and TrBs. We revealed the unique characteristics of AME as the first differentiated cell group emerging from an expanding EPI population<sup>8</sup>. We uncovered specific pathways and transcription factors for TrB subtype separation and functional differences between subtypes and between pre-gastrulation TrBs and fetal TrBs<sup>32</sup>. Unlike mice, human EPI after implantation maintains its transcriptional properties for a steady and prolonged period while acquiring properties for neuron differentiation and vasculature development. Overall, we reveal the molecular and morphogenetic developmental landscape of pre-gastrulation human embryos. These data provide crucial insights into the pluripotency of human pluripotent stem cells and uncover stem-cell self-renewal and differentiation processes, and will inform future strategies to improve in vitro fertilization success rates.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-019-1875-y>.

- Shahbazi, M. N. et al. Self-organization of the human embryo in the absence of maternal tissues. *Nat. Cell Biol.* **18**, 700–708 (2016).
- Deglinert, A. et al. Self-organization of the *in vitro* attached human embryo. *Nature* **533**, 251–254 (2016).
- Rossant, J. Human embryology: implantation barrier overcome. *Nature* **533**, 182–183 (2016).

- Rossant, J. & Tam, P. P. L. New insights into early human development: lessons for stem cell derivation and differentiation. *Cell Stem Cell* **20**, 18–28 (2017).
- Wamaita, S. E. & Nakan, K. K. in *Current Topics in Developmental Biology* Vol. 128 (eds Plusa, B. & Hadjantonakis, A.-K.) 295–338 (Academic, 2018).
- Shao, Y. et al. A pluripotent stem cell-based model for post-implantation human amniotic sac development. *Nat. Commun.* **8**, 208 (2017).
- Martyn, I., Kanno, T. Y., Ruzo, A., Siggia, E. D. & Brivanlou, A. H. Self-organization of a human organizer by combined Wnt and Nodal signalling. *Nature* **558**, 132–135 (2018).
- Shao, Y. et al. Self-organized amniogenesis by human pluripotent stem cells in a biomimetic implantation-like niche. *Nat. Mater.* **16**, 419–425 (2017).
- Watanabe, K. et al. A ROCK inhibitor permits survival of dissociated human embryonic stem cells. *Nat. Biotechnol.* **25**, 681–686 (2007).
- Dobrev, M. P., Pereira, P. N., Deprest, J. & Zwijsen, A. On the origin of amniotic stem cells: of mice and men. *Int. J. Dev. Biol.* **54**, 761–777 (2010).
- Luckett, W. P. Origin and differentiation of the yolk sac and extraembryonic mesoderm in presomite human and rhesus monkey embryos. *Am. J. Anat.* **152**, 59–97 (1978).
- Shahbazi, M. N. et al. Pluripotent state transitions coordinate morphogenesis in mouse and human embryos. *Nature* **552**, 239–243 (2017).
- Enders, A. C., Schlafke, S. & Hendrickx, A. G. Differentiation of the embryonic disc, amnion, and yolk sac in the rhesus monkey. *Am. J. Anat.* **177**, 161–185 (1986).
- Gardner, R. L. Investigation of cell lineage and differentiation in the extraembryonic endoderm of the mouse embryo. *J. Embryol. Exp. Morphol.* **68**, 175–198 (1982).
- Stirpar, G. G. et al. Integrated analysis of single-cell embryo data yields a unified transcriptome signature for the human pre-implantation epiblast. *Development* **145**, dev158501 (2018).
- Blakeley, P. et al. Defining the three cell lineages of the human blastocyst by single-cell RNA-seq. *Development* **142**, 3151–3165 (2015).
- Yan, L. et al. Single-cell RNA-seq profiling of human preimplantation embryos and embryonic stem cells. *Mol. Biol.* **20**, 1131–1139 (2013).
- Petropoulos, S. et al. Single-cell RNA-seq reveals lineage and X chromosome dynamics in human preimplantation embryos. *Cell* **165**, 1012–1026 (2016).
- Boroviak, T. & Nichols, J. Primate embryogenesis predicts the hallmarks of human naive pluripotency. *Development* **144**, 175–186 (2017).
- Halbleib, J. M. & Nelson, W. J. Cadherins in development: cell adhesion, sorting, and tissue morphogenesis. *Genes Dev.* **20**, 3199–3214 (2006).
- Bedzhov, I. & Zernicka-Goetz, M. Self-organizing properties of mouse pluripotent cells initiate morphogenesis upon implantation. *Cell* **156**, 1032–1044 (2014).
- Sozen, B. et al. Self-assembly of embryonic and two extra-embryonic stem cell types into gastrulating embryo-like structures. *Nat. Cell Biol.* **20**, 979–989 (2018).
- Sasaki, K. et al. The germ cell fate of cynomolgus monkeys is specified in the nascent amnion. *Dev. Cell* **39**, 169–185 (2016).
- de Ikonoff, L. K. & Cedard, L. Localization of human chorionic gonadotropin and somatomammotropin hormones by the peroxidase immunohistochemical method in villi and amniotic epithelium of human placentas (from six weeks to term). *Am. J. Obstet. Gynecol.* **116**, 1124–1132 (1973).
- Morgani, S. M., Metzger, J. J., Nichols, J., Siggia, E. D. & Hadjantonakis, A. K. Micropattern differentiation of mouse pluripotent stem cells recapitulates embryo regionalized cell fate patterning. *eLife* **7**, e32839 (2018).
- Tsakiridis, A. et al. Distinct Wnt-driven primitive streak-like populations reflect *in vivo* lineage precursors. *Development* **141**, 1209–1221 (2014).
- Sun, L. T. et al. Nanog co-regulated by Nodal/Smad2 and Oct4 is required for pluripotency in developing mouse epiblast. *Dev. Biol.* **392**, 182–192 (2014).
- Huelsken, J. et al. Requirement for  $\beta$ -catenin in anterior-posterior axis formation in mice. *J. Cell Biol.* **148**, 567–578 (2000).
- Camus, A., Perea-Gomez, A., Moreau, A. & Collignon, J. Absence of Nodal signaling promotes precocious neural differentiation in the mouse embryo. *Dev. Biol.* **295**, 743–755 (2006).
- Voiculescu, O., Bertocchini, F., Wolpert, L., Keller, R. E., Stern, C. D. The amniotic primitive streak is defined by epithelial cell intercalation before gastrulation. *Nature* **449**, 1049–1052 (2007).
- Fong, G.-H., Klingensmith, J., Wood, C. R., Rossant, J. & Breitman, M. L. Regulation of *flt-1* expression during mouse embryogenesis suggests a role in the establishment of vascular endothelium. *Dev. Dyn.* **207**, 1–10 (1996).
- Liu, Y. et al. Single-cell RNA-seq reveals the diversity of trophoblast subtypes and patterns of differentiation in the human placenta. *Cell Res.* **28**, 819–832 (2018).
- Okazaki, H. et al. Derivation of human trophoblast stem cells. *Cell Stem Cell* **22**, 50–63.e56 (2018).
- Vento-Tormo, R. et al. Single-cell reconstruction of the early maternal–fetal interface in humans. *Nature* **563**, 347–353 (2018).
- Smith, A. Formative pluripotency: the executive phase in a developmental continuum. *Development* **144**, 365–373 (2017).
- Hackett, J. A. & Surani, M. A. Regulatory principles of pluripotency: from the ground state up. *Cell Stem Cell* **15**, 416–430 (2014).
- Som, A. et al. The PluriNetWork: an electronic representation of the network underlying pluripotency in mouse, and its applications. *PLoS ONE* **5**, e15165 (2010).
- Liu, D. et al. Single-cell RNA-sequencing reveals the existence of naive and primed pluripotency in pre-implantation rhesus monkey embryos. *Genome Res.* **28**, 1481–1493 (2018).
- Nakamura, T. et al. A developmental coordinate of pluripotency among mice, monkeys and humans. *Nature* **537**, 57–62 (2016).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2019

## Methods

### Ethics statement

This work was approved by the Medicine Ethics Committee of The First People's Hospital of Yunnan Province (2017LS[K]NO.035). All donated embryos in this study were surplus frozen embryos from couples who already had at least one healthy baby after in vitro fertilization clinic treatment. The informed consent process for embryo donation complied with International Society for Stem Cell Research (ISSCR) 'Guidelines for Stem Cell Research and Clinical Translation (2016)' and 'Ethical Guidelines for Human Embryonic Stem Cell Research (2003)' jointly issued by the Ministry of Science and Technology and the Ministry of Health of People's Republic of China. The Medicine Ethics Committee of The First People's Hospital of Yunnan Province is composed of nine members, including lawyers, scientists and clinicians with relevant expertise. The Committee evaluated the scientific merit and ethical justification of this study and conducted a full review of the donations and use of these samples. All donor couples signed informed consents for voluntary donations of surplus embryos for human embryo development study at the Department of Reproductive Medicine in the First People's Hospital of Yunnan Province. No financial inducements were offered for the donations. In the process, couples were informed that their embryos would be used to study the developmental mechanisms of human embryos and that their donation would not affect their in vitro fertilization cycle. The culture of all embryos was terminated at day 14 after fertilization or upon the appearance of primitive streak anlage.

### Embryo thawing and zona pellucida removal

Before embryo thawing, human embryo culture medium G-2 (10132, Vitrolife) was equilibrated in a 4-well plate (176740, Nunc) overnight. First, 0.5 ml G-2 medium was added into the well and 0.25 ml mineral oil (10029, Vitrolife) was used to cover the G-2 medium. Human blastocysts (5–6 d.p.f.) were thawed using Kitazato Thawing Media Kit (VT102, Kitazato Corporation) by following manufacturer's instructions. After culture in drops of the equilibrated G-2 medium for 4 h, embryos were transferred to acidic Tyrode's solution (T1788, Sigma-Aldrich) to remove the zona pellucida. Once the zona pellucida vanished, embryos were immediately transferred to G-2 medium. Embryos were washed in G-2 medium twice and then transferred to the in vitro culture medium.

### Evaluation of embryo quality

According to the Gardner's scoring system<sup>40</sup>, thawed blastocysts were given numerical scores from 1 to 6 based on their expansion degree and hatching status. The blastocyst with expansion and hatching status above 3 and with visible inner cell mass above grade B were included in the study. On the basis of morphologies, healthy embryos had to meet the two following requirements: obvious expansion during culturing and absence of obviously dead or broken (fragmented) cell mass during development. Otherwise, they were excluded from this study.

### In vitro 3D culture of human embryos

Embryos without zona pellucida were cultured in a low attachment 96-well plate (3474, Corning) with 1 embryo and 150 µl blastocyst-culture medium in each well. The embryo culturing conditions were as follows: 37.2 °C, 6% CO<sub>2</sub> and saturated humidity. The culture protocol was summarized in Fig. 1a. First, at 6–8 d.p.f. the culture medium was modified in vitro culture medium 1 (mIVC1). At 8 d.p.f., 50% of mIVC1 medium was replaced by mIVC2. At 9 d.p.f., embryos were transferred to new wells in mIVC2 including 10% Matrigel (354234, Corning). Then, 50% culture medium was replaced by new mIVC2 including 10% Matrigel every day. mIVC1 and mIVC2 were pre-equilibrated in the incubator for at least 6 h before use. A step-by-step protocol has also been included at Protocol Exchange<sup>41</sup>.

The components of the mIVC1 and mIVC2 system are modifications of IVC1 and IVC2, established for culturing human attached embryos

on the dish<sup>1,2</sup>. mIVC1: advanced DMEM/F12 (12634-010, Thermo Fischer Scientific) supplemented with 20% (v/v) DFBS (defined fetal bovine serum) (bs-0003, Biosera), 2 mM L-glutamine (25030, Thermo Fisher Scientific), 1× ITS-X (51500-056, Thermo Fisher Scientific), 8 nM β-oestradiol (E8875, Sigma-Aldrich), 200 ng ml<sup>-1</sup> progesterone (P0130, Sigma-Aldrich), 25 µM *N*-acetyl-L-cysteine (A7250, Sigma-Aldrich), 0.22% (v/v) sodium lactate (L7900, Sigma-Aldrich), 1 mM sodium pyruvate (P4562, Sigma-Aldrich) and 10 µM Y27632 (S1049, Selleck). mIVC2: advanced DMEM/F12 supplemented with 30% (v/v) KOSR (knockout serum replacement) (A3181501, Thermo Fischer Scientific), 2 mM L-glutamine, 1× ITS-X, 8 nM β-oestradiol, 200 ng ml<sup>-1</sup> progesterone, 25 µM *N*-acetyl-L-cysteine, 0.22% (v/v) sodium lactate, 1 mM sodium pyruvate and 10 µM Y27632.

### Optimizing the Matrigel concentration

To identify optimal Matrigel concentration for culturing human blastocysts under 3D condition, we designed the following four group experiments: group 1, W/O Matr: human embryos were cultured on low attachment plate without the Matrigel embedment up to 14 d.p.f.; group 2, 25% Matr: cultured human embryos were embedded in 25% Matrigel at 9 d.p.f.; group 3, Matr+10% Matr: human embryos were embedded in 10% Matrigel and transferred into the new well pre-coated by 100% Matrigel (30 µl) at 9 d.p.f.; group 4, 10% Matr: human embryos were embedded in 10% Matrigel at 9 d.p.f. Culture medium was mIVC1 on 6–7 d.p.f., then 1:1 mIVC1:IVC2 on 8 d.p.f. and mIVC2 on 8–14 d.p.f. For Matr+10% Matr, 30 µl original liquid of Matrigel was added into one well of low attachment 96-well plate at 9 d.p.f. for 30 min in the incubator. Once the Matrigel solidified, 120 µl of mIVC2 with 10% Matrigel was added on the surface of the solidified Matrigel. A single embryo was transferred to a new well after equilibration for 6 h.

### Frozen section staining and taking photographs

Embryos were fixed by 4% paraformaldehyde, washed three times with PBS, dehydrated by 15% sucrose for 1 min and embedded in OCT. Embedded embryos were sectioned by a Leica frozen slicer at a thickness of 10–12 µm. Before staining, the slides were washed by PBS to clear OCT, and permeabilized with 0.2% Triton X-100 for 30 min at room temperature. After blocking with 3% BSA in PBS for 4 h at room temperature, sections were incubated with primary antibodies at 4 °C overnight and then washed three times with 0.05% Tween-20. The following primary antibodies were used: mouse anti-OCT3/4 (Santa Cruz, SC5279, C-10, H1612 1:400), rabbit anti-Brachyury (T) (Santa Cruz, SC20109, polyclonal, A0616, 1:50), rabbit anti-SOX2 (Millipore, AB5603, polyclonal, 2826070, 1:400), goat anti-SOX17 (R&D Systems, AF1924, polyclonal, KGA0815042, 1:250), rabbit anti-KLF4 (Millipore, 09-821, polyclonal, 2383578, 1:400), rabbit anti-KLF17 (Atlas Antibodies, HPA024629, polyclonal, C117502, 1:250), rabbit anti-β-catenin (Abcam, AB32572, E247, GR184212-37, 1:300), mouse anti-E-cadherin (Abcam, AB76055, M168, GR299147-4, 1:100), rabbit anti-cytokeratin 7 (CK7) (Abcam, AB181598, EP17078, GR3214132-10, 1:300), rabbit anti-N-cadherin (Abcam, AB12221, polyclonal, 40975, 1:200), mouse anti-OTX2 (Santa Cruz, SC514195, D-8, G0816, 1:100), goat anti-LEFTY1 (R&D Systems, AF746, polyclonal, CMM0111101, 1:100), goat anti-SOX1 (R&D Systems, AF3369, polyclonal, XUV0618081, 1:500), rabbit anti-PAX6 (Biolegend, 901301, Poly 19013, b267205, 1:500), rabbit anti-laminin (Sigma-Aldrich, L9393, polyclonal, 028M4890V, 1:50), rabbit anti-TEAD4 (Atlas Antibodies, HPA056896, polyclonal, R78063, 1:150), goat anti-FOXA2 (Santa Cruz, SC6554, polyclonal, D1216, 1:100), goat anti-OCT3/4 (Santa Cruz, SC8628, polyclonal, G3201, 1:250), Mouse anti-hCG (Abcam, AB9582, 5H4-E2, GR308272-2, 1:100), rabbit anti-PRDM14 (Millipore, AB4350, polyclonal, 2897240, 1:50), goat anti-TFCP2L1 (R&D Systems, AF5726, polyclonal, CCUG0115021, 1:200), goat anti-NANOG (R&D Systems, AF1997, polyclonal, KJ0514091, 1:250), mouse anti-PODXL (R&D Systems, MAB1658, 222328, JKW0218041, 1:400), mouse anti-HLA-G (Abcam, AB52455, 4H84, GR251679-19, 1:200), goat anti-GATA6



# Article

(R&D Systems, AF1700, polyclonal, KWT0419021, 1:200), mouse anti EZRIN (Sigma-Aldrich, E8897, 3C12, 117M4875V, 1:500), rabbit anti-CER1 (Sigma-Aldrich, HPA019917, polyclonal, R10176, 1:50), phalloidin (F-actin) Alexa Fluor 488 (Thermo Fisher Scientific, A12379, direct labelled, 1749905, 1:300), rabbit anti HESX1 (Abcam, AB246949, polyclonal, GR3267093-1, 1:100), goat anti-Brachyury (T) (R&D Systems, AF2085, polyclonal, KQP0617031, 1:200), goat anti FLK1 (R&D Systems, AF357, polyclonal, Cvisceral endoderm 0617081, 1:100) and WGA Alexa Fluor 647 (Invitrogen, W32466, direct labelled, 1988457, 1:500). The secondary antibodies were incubated at room temperature for 2 h, and the slices were washed three times with 0.05% Tween-20. Pictures were taken by Leica SP8 laser confocal microscope.

## Whole-embryo staining and 3D reconstruction

Fixed embryos were permeabilized by 0.5% Triton X-100 in PBS overnight in a 4 °C refrigerator. Embryos were blocked with 3% BSA in PBS for 4 h at room temperature and then transferred to a new well. Embryos were incubated with primary antibodies for 16–18 h at 4 °C. Embryos were washed three times in PBS including 0.05% Tween-20, and incubated the secondary antibodies for 4 h in room temperature. Embryos were washed three times in PBS including 0.05% Tween-20 and transferred to a well of 8-well IbiTreat  $\mu$ -plates (IB-80826, Ibi GmbH) with an aqueous solution of 60% glycerol aqueous solution to take photographs.

To make 3D videos, a multiphoton microscope Leica TCS SP8 Divisceral endoderm was used. Embryos were mounted in 80% glycerol aqueous solution. Imaging was performed with a Leica TCS SP8 Divisceral endoderm multiphoton microscope with a HC Fluotar VISIR 25 $\times$ /1.00 NA CLARITY-optimized immersion objective with motorized correction ring (Leica Microsystems). All z-stack images were acquired with at a 1,024  $\times$  1,024 pixel resolution and with a z-step of 1  $\mu$ m. Three-dimensional data were deconvoluted with Lightning module of LAS X software (Leica Microsystems). Because of limited penetration ability of multiphoton microscope, we took 200- $\mu$ m-thick pictures in Z-axis.

## Isolation of single cells

Embryos were washed in PBS three times, washed in 0.25% trypsin (T4799; Sigma-Aldrich) twice, incubated with 0.25% trypsin for 15 min at 37 °C and terminated by DFBS. Embryos were dissociated into single cells by repeated pipetting and dispersed in 1% DFBS in PBS. A single cell was pipetted into a PCR tube. All above operations were performed using a Nikon SMZ645 microscopy.

## Cell number counts and embryo diameter measurements

To count cell numbers, two protocols were used. First, the whole embryo was dissociated into single cells by digestion with 0.25% trypsin, and the total cell number per embryo was counted. The second was used to count numbers of specific cell types. After staining, cell numbers of OCT4<sup>+</sup> EPIs, GATA6<sup>+</sup> PrE and CK7<sup>+</sup> TrBs were analysed by ImageJ software (v.1.51j8). To measure the diameter of developing human embryos, embryos were photographed every day and their diameters were measured by the ImageJ software.

## RNA sequencing of single cells

Isolated single cells were washed in DPBS (Gibco 14190-144) and picked up using Pasteur pipettes under a dissecting microscope. The synthesis and amplification of full-length cDNAs were performed following Smart-seq2 protocol<sup>42</sup>. In brief, single cells were washed in DPBS (Gibco 14190-144) and picked into lysis buffer using Pasteur pipettes under a dissecting microscope. Reverse transcription reactions and pre-amplification were performed using SuperScript II (Invitrogen 18064-014) and KAPA HiFi HotStart ReadyMix (KAPA Biosystems KK2601), respectively. The quality of the cDNAs was evaluated by Bioanalyzer 2100. Library construction and sequencing were performed by Annoroad Gene Technology (<http://www.annoroad.com/>) or BGI

(<https://www.bgi.com/>). Sequencing was performed on an Illumina X-ten platform or a BGISEQ-500 sequencing platform (BGI). Pair-end reads were obtained, and the number of the reads was more than 7 million for every individual cell.

## Quality control, alignment of the scRNA-seq profiles and stringent filtering

The sequencing qualities of 557 scRNA-seq profiles were examined with the FASTQC (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) and MULTIQC (v.1.6)<sup>43</sup>. The annotation of RefSeq genes were downloaded from UCSC Genome Browser<sup>44</sup>. RefSeq exons were used to build databases of exon and splice sites with the `extract_exon.py` and `extract_splice_sites.py` in the `hisat2` package<sup>45</sup>. HISAT2 (v.2.1.0)<sup>45</sup> was used to align the scRNA-seq profiles to the human genome. The alignment results of HISAT2 in the SAM format were converted to BAM format and sorted with SAMTools (v.1.1)<sup>46</sup>. Stringtie (v.1.3.4)<sup>47</sup> was used to calculate the abundances of genes (in FPKM) annotated in GENCODE (v.29)<sup>48,49</sup> using the options of ‘-G gencode.v29.gtf -B -e -v’. Because the cells with limited number of expressed genes were potentially caused by RNA degradation, two scRNA-seq profiles with numbers of genes with abundance levels more than 1 FPKM were smaller than 2,000 were eliminated in further analysis. The basic information of the 555 remaining scRNA-seq profiles were available in Supplementary Table 8.

Qualimap 2 (v.2.2.2-dev)<sup>50</sup> was used to calculate the number of reads mapped to the genes in GENCODE genes with options of ‘-java-mem-size=40G comp-counts -bam -pe’. Then, we prepared a plot of the number of genes versus sequencing depth (the number of sequencing reads in the scRNA-seq library) with the command of ‘qualimap counts -d -k 5’.

## The t-SNE and trajectory analysis of the scRNA-seq profiles

Genes with dispersion values of at least 0.5 in a particular cell type were selected with the Seurat package (v.2.3.4) in R<sup>51</sup>. The top variable genes were used to classify cells with the FindClusters function in the Seurat package. Cell types were defined by expression of selected markers. Genes filtered with Seurat were used to perform t-SNE analysis for the cell types under consideration using the RunTSNE function in the Seurat package of R. Monocle (v.2.4.0)<sup>52,53</sup> was used to perform a trajectory analysis for the cell type under consideration. The heat map function of the R platform was used to generate the heat map of selected marker genes. Because AME did not express pluripotency genes (or expressed them only at low levels), we sorted cells from the ICM and EPI clusters by NANOG expression in analysing dynamics of pluripotency and primitive streak genes to exclude the AME and intermediate state cells. One hundred and thirty-six cells with FPKM values of NANOG more than 1 were specifically used for violin plots of pluripotency genes and gene regulatory networks analysis.

## Identifying genes related to different cell types

A feature vector of one cell type was defined as a binary vector with values of 1 for the cell types under consideration and 0 for other cell types. Genes with Pearson’s correlation coefficients of least 0.4 with the feature vector for a particular type of cell were regarded as genes related to the cell type under consideration. One hundred and thirty-six cells with FPKM values of NANOG more than 1 were specifically used for violin plots of pluripotency genes and gene regulatory networks analysis (136 cells reclassified were used in Extended Data Figs. 5g, 19e, f–i).

## Comparisons with publicly available scRNA-seq from pre-implantation human embryos

The previously described analytical strategies and datasets<sup>15</sup> were used to combine and analyse our scRNA-seq data from 6–9-d.p.f. embryos and the data (later blastocyst or 6–7-d.p.f. blastocysts) from three previous studies with PCA<sup>16–18</sup>. The 12 previously described lineage marker genes<sup>15</sup> were used in the PCA analysis. The combined dataset

was aligned to genome to calculate the FPKM values of genes, and analysed by Seurat (v.2.3.4).

### Identification of differentially expressed genes between groups

Differentially expressed genes between two groups were obtained by using edgeR<sup>54</sup>. Genes with uncorrected *P* values (likelihood ratio tests) smaller than 0.01, absolute fold difference larger than 2 and median of FPKM larger than 1 in one group were regarded as differentially expressed genes.

For HESX1<sup>+</sup>T<sup>-</sup> and HESX1<sup>-</sup>T<sup>+</sup> cells, EPIs at 14 d.p.f. were classified as two groups, HESX1<sup>+</sup>T<sup>-</sup> and HESX1<sup>-</sup>T<sup>+</sup>, based on HESX1 and T expression in these cells. Cells positive for *T* (FPKM > 1) and negative for *HESX1* (FPKM < 1) belonged to the HESX1<sup>-</sup>T<sup>+</sup> group, and cells negative for *T* (FPKM < 1) and positive for *HESX1* (FPKM > 1) belonged to HESX1<sup>+</sup>T<sup>-</sup> group. The lists of down- and upregulated genes in HESX1<sup>+</sup>T<sup>-</sup> cells compared to HESX1<sup>-</sup>T<sup>+</sup> cells are shown in Supplementary Table 3.

### Co-expression gene network analysis

In total, 181 scRNA-seq profiles of 7–9-d.p.f. blastocysts (in Supplementary Table 1) were used to perform co-expression gene networks (11 cells were not included owing to high expression of two lineage-marker genes). These 181 cells were assigned as EPI, PrE and TrB cells based on expression of selected marker genes. We filtered the genes by keeping those with log<sub>2</sub>-scaled expression (FPKM + 1) values larger than 4 in at least one sample and with correlation coefficients of at least 0.4 with the feature vector of one of the three cell types (that is, EPI, PrE and TrB). Filtered genes were used to construct co-expression networks of genes with WGCNA<sup>55</sup>. Three gene modules with the highest correlation coefficient values in the three cell types were used to draw gene modules for each cell type. Only transcription factors in each module were kept when visualizing the three gene modules with Cytoscape (v.3.6.1)<sup>56,57</sup>. The eigengene value matrix between each scRNA-seq profile and the nine identified modules calculated by WGCNA analysis was used to perform a 2D hierarchical clustering and visualized with pheatmap function of the R platform.

### GO and KEGG pathway analysis

Enriched GO terms and KEGG pathways for the genes related to different types of cells were identified using KOBAS3<sup>58</sup>. Significant GO terms and KEGG pathways were visualized with the ggplot function in the ggplot2 package in R.

### Gene regulatory network analysis

The PluriNetWork<sup>37</sup> was trimmed by keeping genes and their relations, if genes had at least 10 FPKM in 60% of the subtype of EPI cells or if the gene expression levels were three times higher in the subtype of EPI cells than other cells. The trimmed networks were visualized with Cytoscape (v.3.6.1)<sup>55,56</sup>.

### Comparisons of human and monkey EPI development by analysing scRNA-seq profiles

The genome and annotation of cynomolgus monkey (version mfa5.0) were downloaded from the NCBI Genome database. scRNA-seq profiles of 213 cynomolgus monkey (*Macaca fascicularis*) EPI cells reported previously<sup>39</sup> were aligned to the monkey genome with the same options when analysing human scRNA-seq profiles. The 222 human EPI and 213 cynomolgus monkey EPI scRNA-seq profiles were combined to keep the FPKM values of the 16,487 homologous genes. The 16,487 common genes of human and cynomolgus monkey were filtered to keep 12,475 genes with log<sub>2</sub>(FPKM + 1) values of at least 4 in at least one of the 222 human or 213 cynomolgus monkey cells. Raw FPKM values of the 12,475 genes for the 222 human or 213 cynomolgus monkey cells were normalized with the ‘normalize data’ and scaled with the ‘scale data’ of the Seurat package in R<sup>51</sup>. Normalized and scaled FPKM values were used to perform PCA analysis by using the prcomp function in

R. PCA results were visualized with MatLab (MathWorks) for the first three principle components. The second and third components were also used to visualize the PCA results with the ggplot function in the ggplot2 package of R. A total of 966 homologous genes in human and monkey that contributed highly to PC1 (with absolute PC1 values of more than 2 s.d. of the 12,475 genes) were clustered and visualized with the heatmap function of the R platform. The 294 and 672 genes with PC1 values of >2 and <-2 s.d. were used to perform GO enrichment analysis with KOBAS3<sup>58</sup>. In total, 1,152 genes with significant loading scores for PC2 and PC3 (radius of PC2 and PC3 > 3 s.d and none of the 1,151 gene overlapping with 996 genes that had significant scores for PC1 loading) were clustered and visualized with the pheatmap function of the R platform.

### Statistical analysis

No statistical methods were used to predetermine sample size. The experiments were not randomized, and investigators were not blinded to allocation during experiments and outcome assessment. Errors and error bars represent s.e.m. from a minimum of five independent embryos unless otherwise indicated. Figures display representative results. Unless otherwise specified, the results were the same across all the embryos analysed. For cell number, the significance difference between two samples was evaluated by unpaired two-sample Student's *t*-test using Excel software (2016). For gene expression, the differences in different cell types were analysed by Wilcoxon rank-sum test. *P* < 0.05 was considered as statistically significant differences.

### Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

### Data availability

Source Data for Extended Data Figs. 1f, k–m, 4b, f, p, 5g, h, l, m, p, q, r, 9f, 10d–h are provided with the paper. scRNA-seq data have been deposited in the Gene Expression Omnibus (GEO) under accession number GSE136447 (scRNA-seq data). The SC3-seq data of cynomolgus monkey embryos (for Extended Data Fig. 10) are GEO accession GSE74767 (ref. <sup>39</sup>); the scRNA-seq data of human pre-implantation embryos (for Extended Data Fig. 3d, e) are with GEO accession GSE66507 (ref. <sup>16</sup>) and GSE36552 (ref. <sup>17</sup>) and ArrayExpress accession E-MTAB-3929 (ref. <sup>18</sup>).

40. Gardner, D. K., Lane, M., Stevens, J., Schlenker, T. & Schoolcraft, W. B. Blastocyst score affects implantation and pregnancy outcome: towards a single blastocyst transfer. *Fertil. Steril.* **73**, 1155–1158 (2000).
41. Xiang, L., Yin, Y. & Li, T. Protocol for a developmental landscape of 3D-cultured human pre-gastrulation embryos. *Protoc. Exch.* <https://doi.org/10.21203/rs.2.16169/v1> (2019).
42. Picelli, S. et al. Full-length RNA-seq from single cells using Smart-seq2. *Nat. Protocols* **9**, 171–181 (2014).
43. Ewels, P., Magnusson, M., Lundin, S. & Källér, M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047–3048 (2016).
44. Rosenbloom, K. R. et al. The UCSC Genome Browser database: 2015 update. *Nucleic Acids Res.* **43**, D670–D681 (2015).
45. Kim, D., Langmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* **12**, 357–360 (2015).
46. Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
47. Perte, M. et al. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* **33**, 290–295 (2015).
48. Harrow, J. et al. GENCODE: the reference human genome annotation for The ENCODE Project. *Genome Res.* **22**, 1760–1774 (2012).
49. Frankish, A. et al. GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res.* **47**, D766–D773 (2019).
50. Okonechnikov, K., Conesa, A. & García-Alcalde, F. Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics* **32**, 292–294 (2016).
51. Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420 (2018).
52. Trapnell, C. et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* **32**, 381–386 (2014).
53. Qiu, X. et al. Reversed graph embedding resolves complex single-cell trajectories. *Nat. Methods* **14**, 979–982 (2017).

54. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
55. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559 (2008).
56. Shannon, P. et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
57. Smoot, M. E., Ono, K., Ruscheinski, J., Wang, P. L. & Ideker, T. Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* **27**, 431–432 (2011).
58. Xie, C. et al. KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Res.* **39**, W316–22 (2011).

**Acknowledgements** This work was supported by the National Key Research and Development Program of China (2018YFA0108500), the Yunnan Key R&D Program (2018ZF007-02 and 2018FA008), the National Natural Science Foundation of China (31760268 and 81660266), Chinese Postdoctoral Science Foundation (2017M623317XB),

Yunnan Provincial Innovation Team (2018HC003 and 2017HC009) and Yunnan Innovation Talents of Science and Technology (2013HA027). We thank Brandi J. Mattson for editing assistance.

**Author contributions** T.L. and W.J. initiated the project. T.L. designed the experiments, organized and supervised the entire project, and wrote the manuscript. T.L. supervised L.X. and Y.Y. to perform data analysis. L.X., Y.Y., Y.M., Y.L., Z.A., D.W., K.D., J. H. and Y.B. performed embryo culture, single-cell isolation and data collection. Y. Z., Y.Y., Z.Z., J.G., S.R. Z.S., X.D. and Y.N. analysed RNA-seq data.

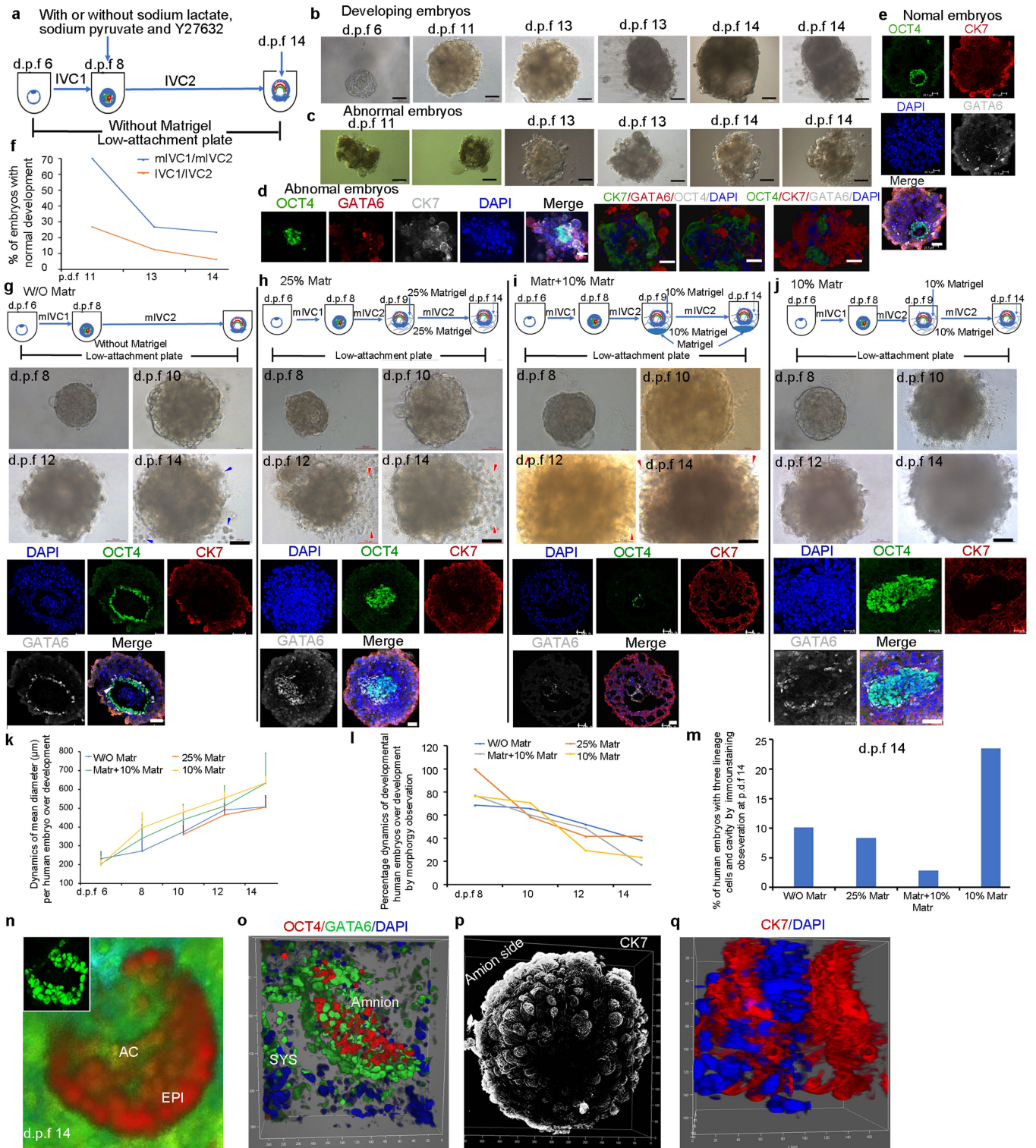
**Competing interests** The authors declare no competing interests.

#### Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41586-019-1875-y>.

**Correspondence and requests for materials** should be addressed to W.J. or T.L.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>.

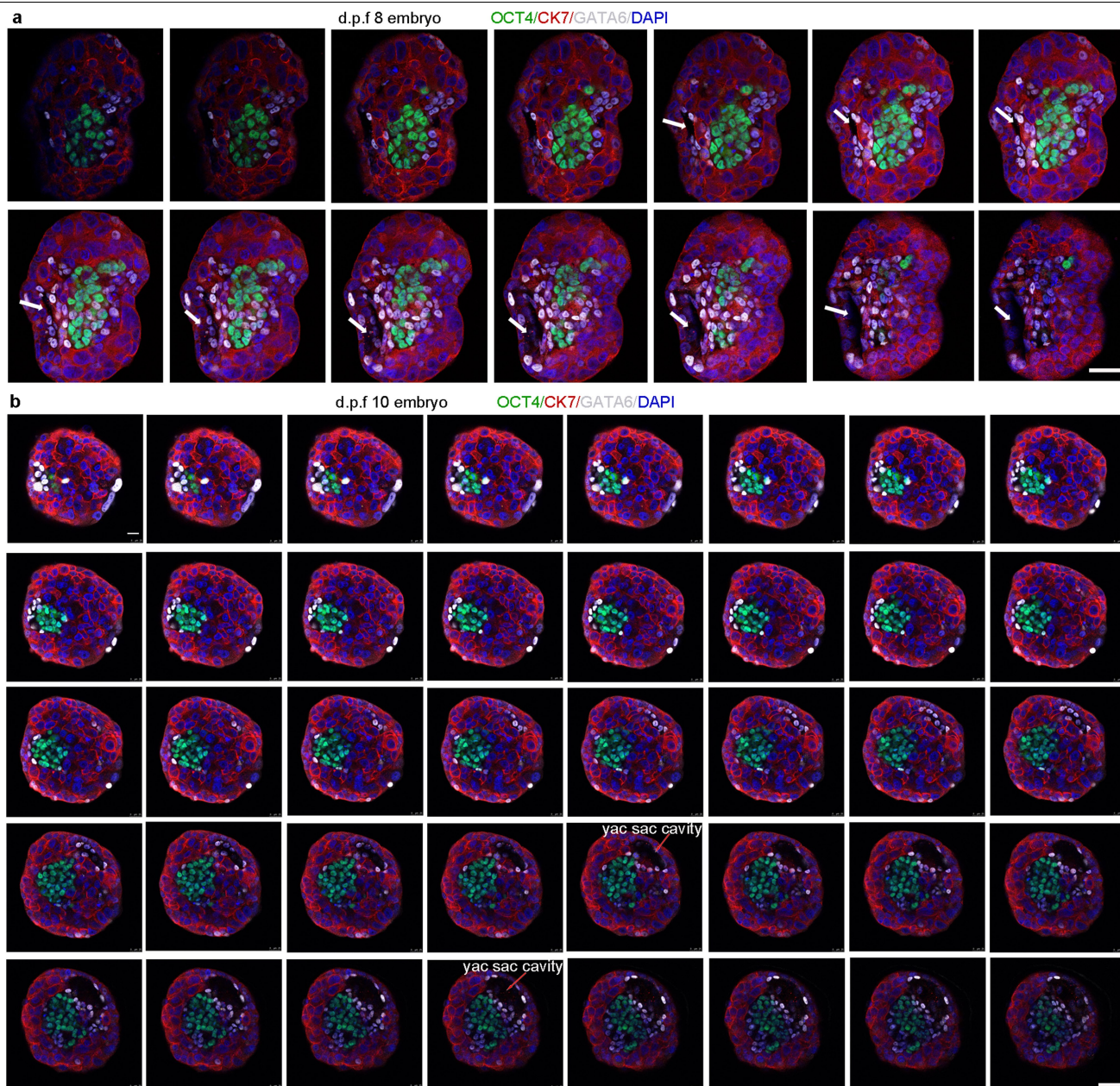


**Extended Data Fig. 1** | See next page for caption.



**Extended Data Fig. 1 | Establishing a 3D blastocyst-culture system.** This figure is related to Fig. 1. **a–f**, Modification of human embryo culture medium. To test human embryo development, human embryos at 5–6 d.p.f. were cultured on low attachment plate without any 3D extracellular matrix up to 14 d.p.f. By sequentially culturing in IVC1 (6–8 d.p.f.) and IVC2 (8–14 d.p.f.) media, 6.67% (2 out of 30 embryos) of embryos could survive until 14 d.p.f. **a**, Schematics of improving culture medium. Sodium lactate, sodium pyruvate and ROCK inhibitor (Y27632) were added to the IVC1 and IVC2 media, referred to as mIVC1 (6–8 d.p.f.) and mIVC2 (8–14 d.p.f.), respectively. Culture in mIVC1 and mIVC2 enabled 25% (4 out of 16 embryos) of human blastocysts to develop up to 14 d.p.f. **(c)**, **b**, Representative developing embryos based on morphological observation.  $n = 16$  independent embryos from three independent experiments. **c**, Representative embryos with abnormal development. Abnormal embryos displayed growth arrest or had obvious cell death or fragmentations.  $n = 13$  independent embryos from three independent experiments. **d**, Representative staining of abnormal embryos with CK7, GATA6 and OCT4 at 14 d.p.f. In all six examined embryos, consistent data were obtained. **e**, Staining of developing embryos (2 out of 3 embryos) with CK7, GATA6 and OCT4 at 14 d.p.f. **f**, Quantification of developmental rates of human embryos cultured in control medium (IVC1 and IVC2) and modified medium (mIVC1 and mIVC2).  $n = 30$  and 16 biologically independent embryos, respectively. Developmental rates of human embryos at 11 and 13 d.p.f. were based on the two following requirements by morphology: obvious expansion over culture and absence of obvious cell death mass and fragmented phenotypes. At 14 d.p.f., we determined the embryo development ratio by staining CK7 (TrB), GATA6 (PrE) and OCT4 (EPI). **g–q**, Representative human embryo development after culture under four different 3D conditions over development. The limited number of embryos only enabled us to compare embryo development under four conditions. As the implantation time window is 8–10 d.p.f., we embedded embryos with Matrigel at 9 d.p.f. Embryo development was verified on basis of morphological observation and staining of specific markers for OCT4, GATA6 and CK7. **g–j**, Top, schematics of in vitro 3D culture of human blastocysts under different culture conditions. Middle, representative images of human embryos during development. Bottom, representative stained images of cultured human embryos under different 3D conditions at 14 d.p.f. **g**, Human embryos at 5–6 d.p.f. were cultured on low attachment plate without Matrigel (W/O Matr) up to 14 d.p.f. The outermost TrBs showed signs of apoptosis (blue arrowheads), as determined by morphological observations and CK7 staining, which suggests that the condition was unsuitable for TrB development and survival required for attachment. In total, 2 of 20 embryos (three independent experiments) survived up to 14 d.p.f. and displayed normal EPI, PrE and TrB development. **h**, Human embryos were embedded in 25% Matrigel at 9 d.p.f. for continuous

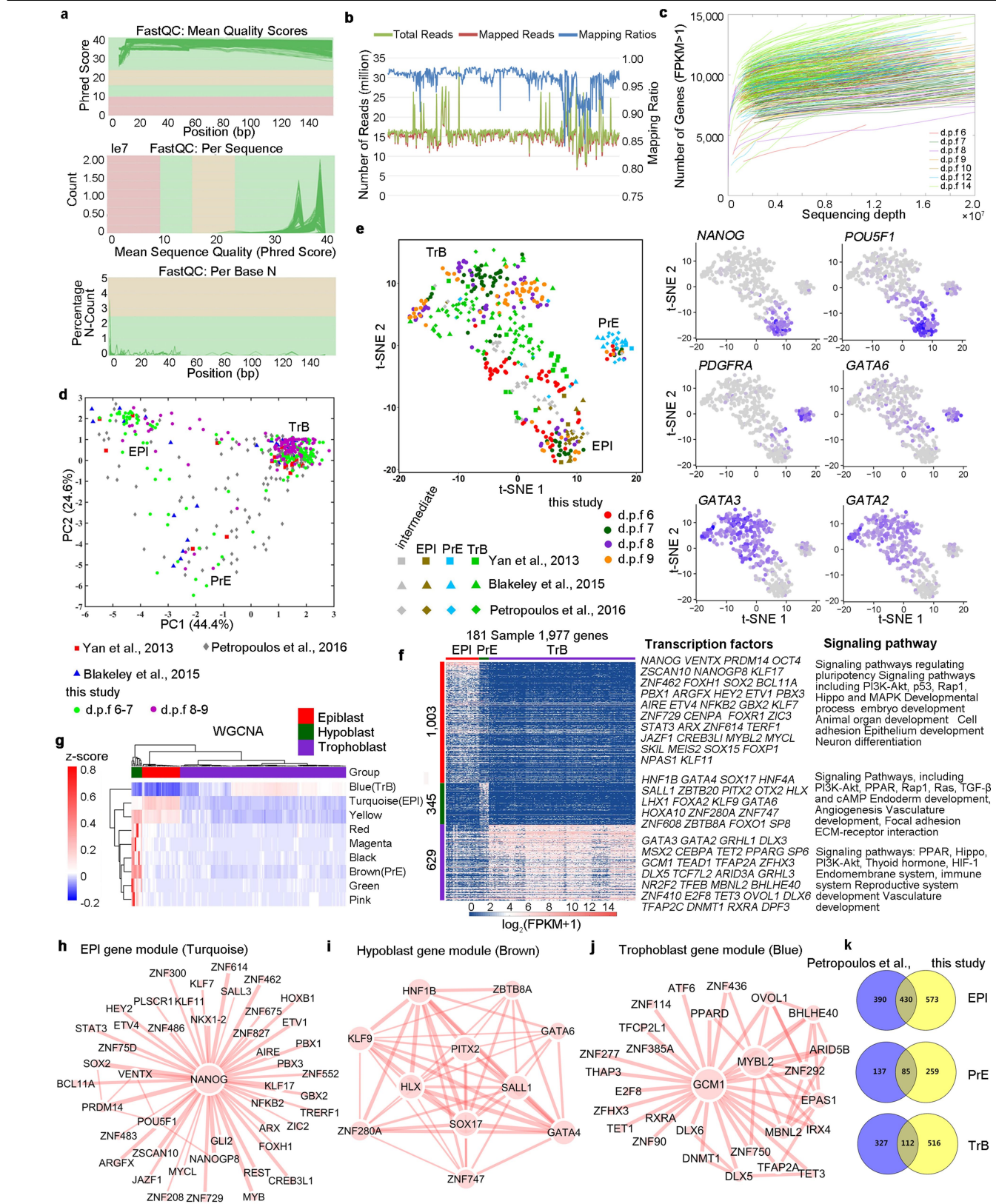
culture up to 14 d.p.f.  $n = 25$  embryos from three independent experiments. The invasion and outgrowth of TrBs were observed and embryos became flat, which suggests a higher concentration of Matrigel is advantageous to differentiation and development of TrBs, as confirmed by CK7 staining. **i**, Human embryos were embedded in 10% Matrigel on the new well, which was pre-coated with 100% Matrigel (30  $\mu$ l) (Matr+10%Matr), at 9 d.p.f. for continuous culture up to 14 d.p.f. Although embryos displayed considerable expansion over culture, staining with lineage markers showed that EPIs in most embryos were lost over development. Only 1 of 33 embryos from three independent experiments grew to 14 d.p.f. and was accompanied by EPI, PrE and TrB development. The negative outcome may indicate that high concentrations of Matrigel can inhibit EPI development. In **h** and **i**, red arrowheads indicate TrBs invading into Matrigel. **j**, Human embryos were embedded in 10% Matrigel at 9 d.p.f. for continuous culture up to 14 d.p.f. Compared with the 25% Matrigel and Matr+10% Matr conditions, human embryos in 10% Matrigel increased in size at the thickness (Z-axis) and showed better 3D spatial structures. In total, 4 of 17 embryos from three independent experiments grew to 14 d.p.f. with normal development. **k**, Quantification of the mean diameter of human embryos cultured under different 3D conditions during development by analysis of 5–14 embryos from three independent experiments. Data are mean  $\pm$  s.d. **l**, Quantification of human developmental embryos during culture in different 3D conditions. Data were based on morphological observations only. Human developing embryos met the two following requirements: obvious expansion over culture; absence of obvious cell death mass or fragmented phenotypes. **m**, Quantification of developing embryos in different 3D conditions.  $n = 20$  (W/O Matr), 25 (25% Matr), 33 (Matr+10% Matr) and 17 (10% Matr) blastocysts. Developing embryos met the following requirements: obvious expansion over culture; absence of obvious cell death mass or fragmented phenotypes; and development of EPIs, PrEs and TrBs and formation of amnion identified by OCT4, GATA6 and CK7 staining. Although embryos in the 25% Matr and Matr+10% Matr culture conditions have normal morphologies, some embryos lacked OCT4<sup>+</sup> EPIs or GATA6<sup>+</sup> PrEs and gave rise to a high proportion of TrBs, which suggest that high concentrations of Matrigel could inhibit EPI development and promote TrB proliferation. **n–q**, Three-dimensional construction of human embryos cultured in 10% Matrigel (3 out of 3 embryos from three independent experiments). **n**, A representative 3D reconstruction of EPIs. Inset shows OCT4 staining of one section from the same embryo. **o**, A representative 3D reconstruction of SYS and amnion including an embryonic disc and an amniotic cavity (see Supplementary Video 1). **p**, Three-dimensional reconstruction of TrBs (see Supplementary Video 2). **q**, Three-dimensional magnification of TrBs close to the amnion side. Scale bars, 100  $\mu$ m (phase-contrast) or 50  $\mu$ m (staining).



**Extended Data Fig. 2 | Representative z-series of an 8- and a 10-d.p.f. embryo.**

This figure is related to Fig. 1f, g. **a**, Series of confocal z-sections of the 8-d.p.f. embryo stained for OCT4 (green), GATA6 (grey) and CK7 (red). The PYS cavity (white arrows) was surrounded by few GATA6<sup>+</sup> PrEs. Similar phenotypes were observed in 3 out of 4 embryos from three independent experiments. **b**, Series

of confocal z-sections of the whole embryo stained for OCT4 (green), GATA6 (grey) and CK7 (red) showing that the PYS cavity in the 10-d.p.f. embryo becomes more distinct. Similar phenotypes were observed in 3 out of 3 embryos from three independent experiments. The yolk sac cavity (red arrows) was surrounded by GATA6<sup>+</sup> PrEs. Scale bars, 50 μm.



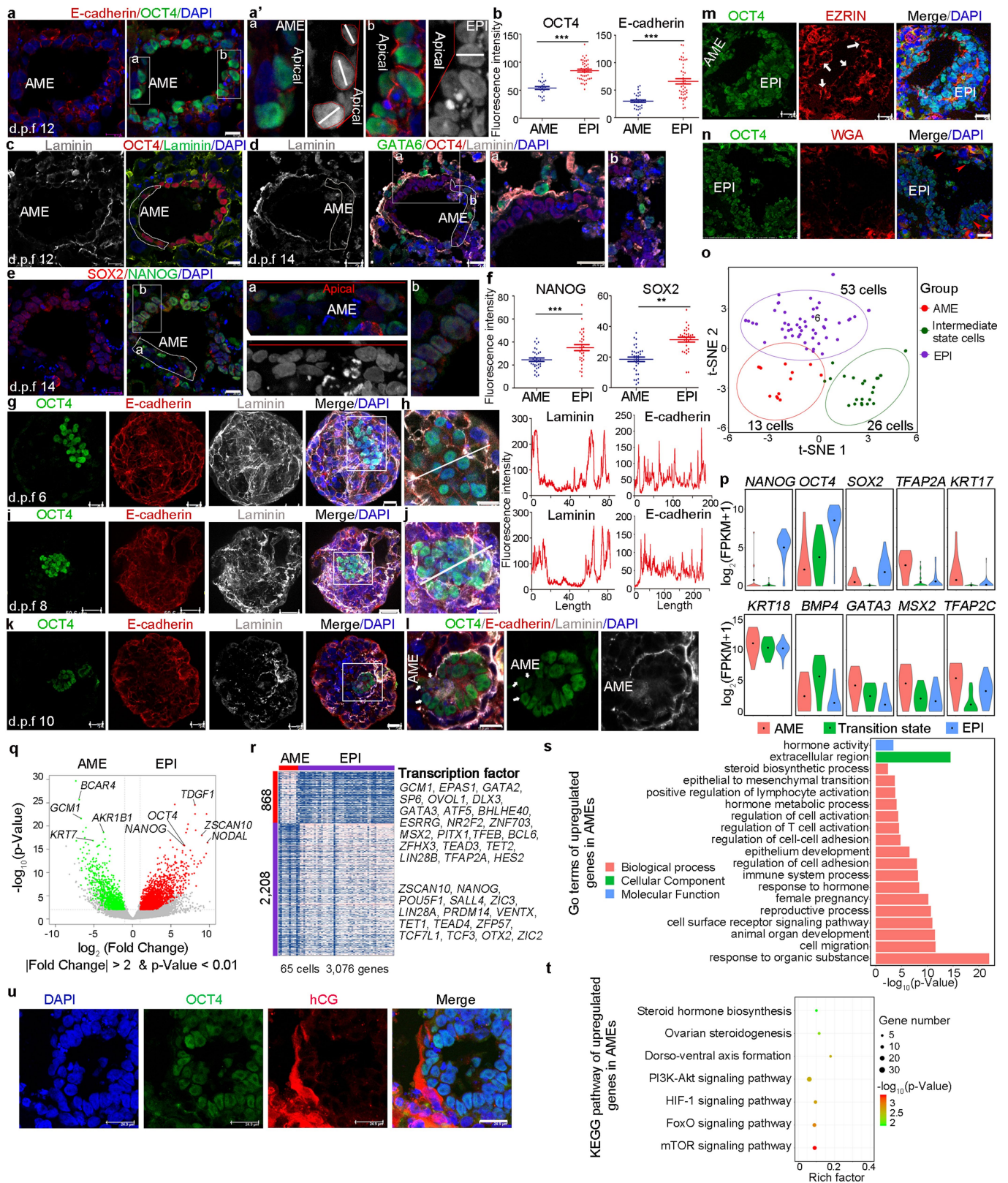
**Extended Data Fig. 3** | See next page for caption.



**Extended Data Fig. 3 | Lineage delineation by transcriptome.** This figure is related to Fig. 2. **a–c**, Quality control of single-cell RNA-sequencing data. **a**, Sequence quality was evaluated by FastQC. **b**, Total reads, mapped reads and mapping ratios of 557 single cells. **c**, Saturation curve of sequencing. **d, e**, Integrated analysis of embryonic single-cell data from different source. We used the analytical strategies developed previously<sup>15</sup> to analyse single-cell RNA-seq data (255 single cells) from 6–9-d.p.f. embryos in the study and single-cell RNA-seq data (216 single cells) from three previous reports<sup>16–18</sup> (later blastocysts or 6–7-d.p.f. blastocysts). The three datasets have previously been analysed<sup>15</sup>. **d**, PCA based on 12 lineage markers (*NANOG*, *SOX2*, *KLF17* and *TDGF1* for EPI; *PDGFRA*, *GATA6*, *GATA4* and *SOX17* for PrE; *GATA3*, *GATA2*, *KRT18* and *TEAD3* for TrB) showed clear separation between EPI, TrB or PrE could be attained for nearly all samples including our single cells from 6–9-d.p.f. blastocysts, which indicates that lineage delamination occurs at 6 d.p.f. The result is consistent with previous findings<sup>15</sup>. **e**, *t*-SNE analyses using 4,333 viable genes across all samples. The samples from previous studies were defined into four types: intermediate cells, EPI, PrE and TrB<sup>15</sup>. The combined Seurat revealed that most of cells independent of cell resource mixed well. Although most of samples were clustered into EPI, PrE or TrB, similar to the results using 12 lineage genes (**d**), some cells from 6-d.p.f. embryos remained in an intermediate state with overlapping expression of *POUSF1*, *GATA6*, *PDGFRA* and *GATA3*. Compared with cells from 6-d.p.f. embryos, cells from 7–9-d.p.f. embryos displayed a clearer separation. These data showed that cell fates of 7–9-d.p.f. embryos became more fixed. **f–j**, Lineage delineation by transcriptome. Analysis of genes corresponding to EPI, PrE and TrB from 7–9-d.p.f. embryos to understand the regulators involved in the segregation process. **f**, Heat map of lineage-specific genes of EPI, TrB and PrE from 7–9-d.p.f. embryos (Supplementary Table 1). Their representative transcriptional factors and KEGG pathways are shown, respectively. GO terms and KEGG pathways showed EPI-specific genes associated with signalling pathways regulating stem-cell pluripotency including PI3K–AKT, p53, RAP1 and MAPK.

PrE-expressing genes related to TGF $\beta$ , PPAR and Ras signalling pathways. TrB-specific genes contributed to Hippo, HIF, PPAR and thyroid receptor signalling pathways. Notably, the PI3K–AKT signalling pathway was enriched in EPIs, PrEs and TrBs. To explain the difference, we examined gene expressions of the PI3K–AKT signalling pathway components in three cell types and found that cell types specifically expressed different genes of PI3K–AKT pathway (Supplementary Table 1.4). **g**, WGCNA dendrogram indicating different gene modules in all single cell samples from 7–9-d.p.f. embryos. Three major branches corresponded to PrE (brown module), TrB (blue module) and EPI (turquoise module). **h–j**, Hub-gene-network analysis of transcriptional factors specific for PrE (brown module), TrB (blue module) and EPIs (turquoise module). The size of dots represents hubness. **h**, Hub-gene network of the EPI-specific gene module. In addition to well-known transcription factors (*NANOG*, *PRDM14*, *SOX2*, *OCT4* (also known as *POUSF1*), *ZSCAN10* and *KLF17*), new candidate factors may associate with EPI differentiation, such as *VENTX*, *BCL11A*, *PBX1* and *ARGFX*. **i**, Hub-gene network of the PrE-specific gene module. High correlations of some transcription factors with PrE differentiation included *GATA4*, *SOX17*, *GATA6* and *HNF1B*. **j**, Hub-gene network of the TrB-specific gene module. TrB-specific transcription factors, such as *MYBL2*, *TFAP2A*, *DLX6* and *GCM1*. **k**, Comparison of lineage-specific total genes overlapping between a previous study<sup>18</sup> and this study. In the previous study<sup>18</sup>, by analysing 5–7-d.p.f. embryos and combining the lineage-specific results, 439, 820 and 222 genes—which significantly maintained TrB-, EPI- and PrE-specific genes, respectively—were identified. Comparison analysis showed that although we identified more EPI-, PrE- and TrB-specific genes by our resource data, core lineage transcription factors (*NANOG*, *POUSF1* and *SOX2* for EPI; *GATA6*, *SOX17* and *GATA4* for PrE; *GATA2* and *GATA3* for TrB) are maintained across different samples. The difference in gene expressions may be contributed by different development stages of embryos. Difference of gene expressions, including transcription factors, is summarized in Supplementary Table 1.6.

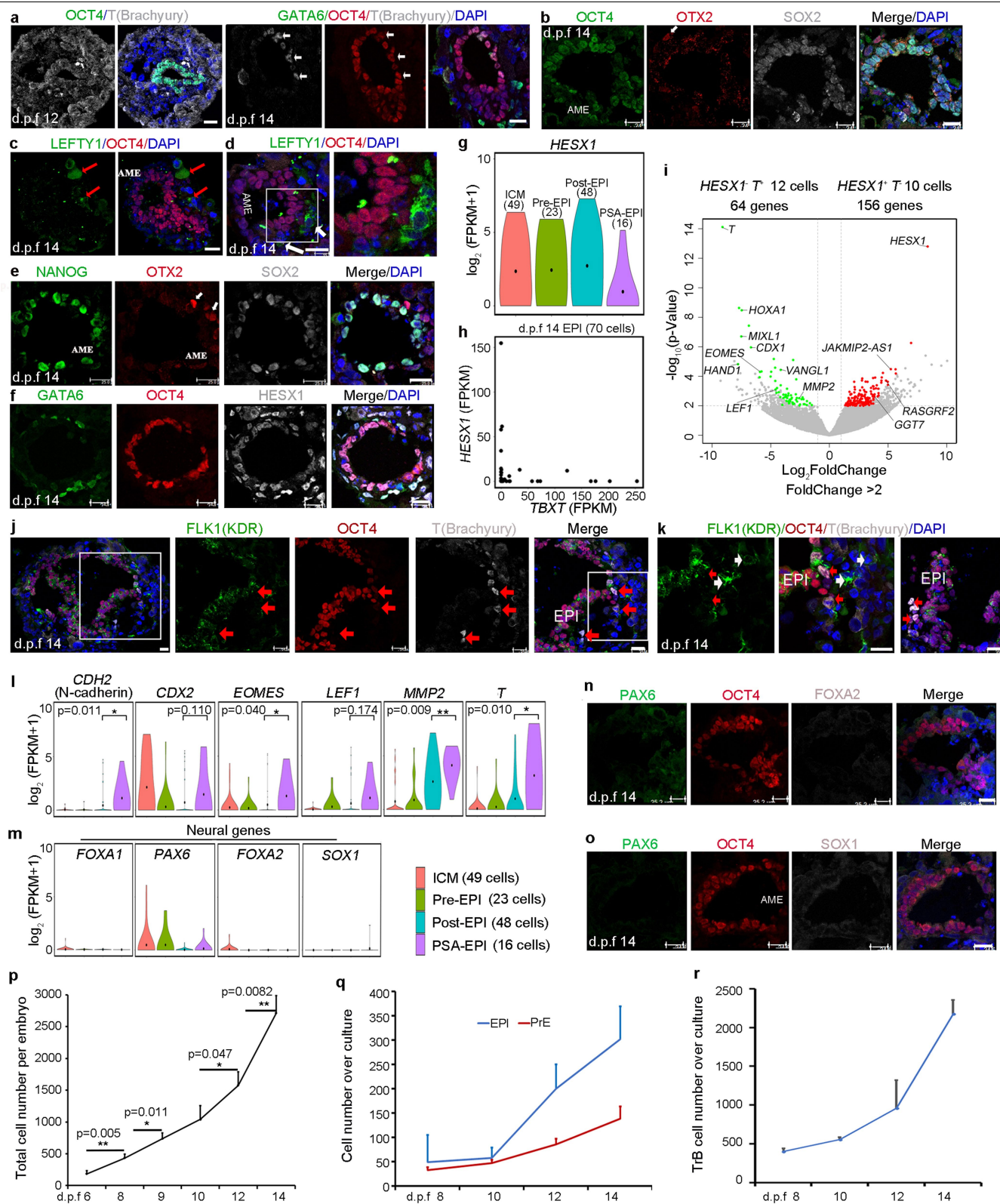




Extended Data Fig. 4 | See next page for caption.

**Extended Data Fig. 4 | AME separation from EPI.** **a**, E-cadherin expression in amnion from a 12-d.p.f. embryo. Right panels show magnified squares (3 out of 3 embryos). White and red lines indicate nuclei and apical orientation, respectively. **b**, Quantification of expression of E-cadherin (AME = 33 cells, EPI = 44 cells) and OCT4 (AME = 22 cells, EPI = 44 cells) in columnar EPIs and squamous AME. Data are mean  $\pm$  s.d from 3 embryos. \*\*\* $P < 0.001$ , two-sided  $t$ -test. **c, d**, Representative staining of laminin at 12 and 14 d.p.f. (3 out of 3 embryos). Right panels show magnified squares from **d**. **e, f**, Representative staining and quantification of NANOG expression (AME = 33 cells, EPI = 30 cells, and SOX2 expression (AME = 34 cells, EPI = 30 cells) in 14-d.p.f. EPIs and AME (3 out of 3 embryos). Data are mean  $\pm$  s.d. \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , two-tailed  $t$ -test. Right panels show magnified squares from **e**. **g, h**, Representative staining and quantification of laminin and E-cadherin in 6-d.p.f. embryos (3 out of 3 embryos). **i, j**, Representative staining and quantification of laminin and E-cadherin in 8-d.p.f. embryos (3 embryos). White long lines in **h** and **j** show positions used to plot intensity profiles (right). **k, l**, Representative staining of laminin and E-cadherin in 10-d.p.f. embryos (3 out of 3 embryos). Together, we conclude that AME separation from EPIs correlates with asymmetrical distributions of E-cadherin and laminin (**a-l**). **m**, Representative staining of EZRIN in 14-d.p.f. amnion (2 out of 2 embryos). White arrows indicate EZRIN expression in apical surface in EPIs and AME. **n**, Representative staining of WGA

in 14-d.p.f. embryos (2 out of 2 embryos). Red arrowheads indicate WGA expression in extra-embryonic cells. **o**,  $t$ -SNE analyses revealed three clusters of 12- and 14-d.p.f. embryos—AME, intermediate state cells and EPIs. **p**, Compared to EPIs, the violin plots show AME significantly downregulated pluripotency genes and upregulated genes that specifically expressed in the AME of 12–17-d.p.f. monkey embryos or self-organized amnion from human pluripotent stem cells. All violins have the same maximum width; black dot denotes the mean. **o, p**, AME,  $n = 13$  cells; intermediate state cells,  $n = 26$  cells; EPI, 53 cells. **q–t**, Gene expression profiles of AME and EPI in the 12- and 14-d.p.f. embryos (Supplementary Table 2). AME,  $n = 12$  cells (one single cell with high *NANOG* expression was not included); EPI, 53 cells. **q**, Volcano map of differentially expressed genes (DEGs) between AME and EPI in the 12- and 14-d.p.f. embryos. DEGs were defined with uncorrected  $P < 0.01$  (two-sided likelihood ratio tests) and  $\log_2$ -transformed fold change  $> 1$  or  $< -1$ , and median FPKM  $> 1$  in one group. **r**, Heat map of DEGs between the AME and EPI. Right panel presents representative transcription factors. **s**, Compared to the EPIs, GO terms of upregulated genes in the AME. **t**, Compared to EPI, KEGG pathways of genes enriched in the AME. **u**, hCG was expressed in the AMEs, but not in the EPIs (2 out of 2 embryos). White arrows indicate the AMEs have squamous nuclear shape, expressed hCG, but downregulated the pluripotent gene, *OCT4*. Scale bars, 15  $\mu\text{m}$  (**a, l**), 20  $\mu\text{m}$  (**c, e**), 25  $\mu\text{m}$  (**d, g, h, k, m, n, u**) or 50  $\mu\text{m}$  (**i**).



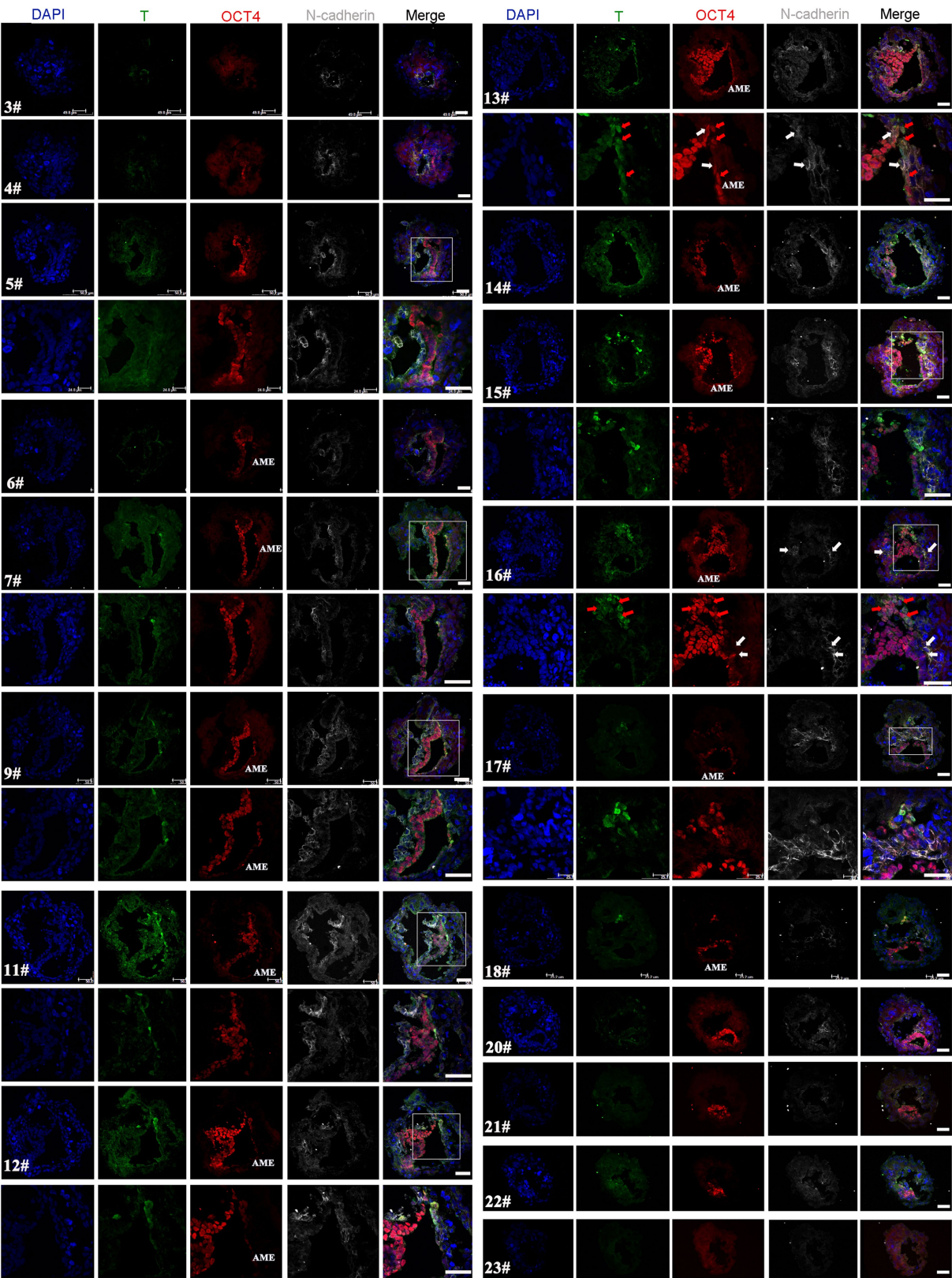
Extended Data Fig. 5 | See next page for caption.



**Extended Data Fig. 5 | Human embryos at 14 d.p.f. initiate anterior-posterior polarity and generation of PSA in 3D-culture conditions.** This figure relates to Fig. 3. **a**, OCT4, GATA6 and TBXT T staining of sections from a 14-d.p.f. embryo (2 out of 2 embryos), showing that T<sup>+</sup> cells originated from the EPI compartment close to the AME compartment boundary at 14 d.p.f. **b**, Representative OCT4, OTX2 and SOX2 staining. Arrow indicates an OTX2<sup>+</sup> cell. *n* = 4 of 5 embryos from three independent experiments displayed consistent data. **c–e**, LEFTY1 (**c**, **d**; *n* = 3 out of 4 embryos from two independent experiments) and OTX2 (**e**; *n* = 3 out of 5 embryos from two independent experiments) immunofluorescence was only detected on the side of 14-d.p.f. embryonic disc. Arrows indicate LEFTY1<sup>+</sup> or OTX2<sup>+</sup> cells. Right image in **d** is magnification of the square in the left image. The exclusive expression of NANOG and SOX2 was not observed in EPIs (**e**). **f**, Staining of OCT4, HESX1 and GATA6 in the 12-d.p.f. embryos (2 out of 2 embryos). **g**, The violin plots show dynamic expression of *HESX1* during EPI development. All violin plots have the same maximum width, black dot denotes the mean. **h**, Correlation of *HESX1* and *T* expression of 14-d.p.f. EPIs, as determined by scRNA-seq. Each plot represents a single cell. **i**, Volcano plots show DEGs in HESX1<sup>+</sup>T<sup>+</sup> (10 single cells) and HESX1<sup>+</sup>T<sup>−</sup> (12 single cells) EPIs by scRNA-seq. DEGs were defined as those with uncorrected *P* < 0.01 (likelihood ratio test) and fold change of >2 or <−2, and median FPKM > 1 in one group. **j**, **k**, Staining of OCT4, FLK1 and T at 14 d.p.f. (2 out of 2 embryos from two independent experiments). Red arrows denote migrating T<sup>+</sup> cells; white arrows denote FLK1<sup>+</sup> extra-embryonic mesenchyme.

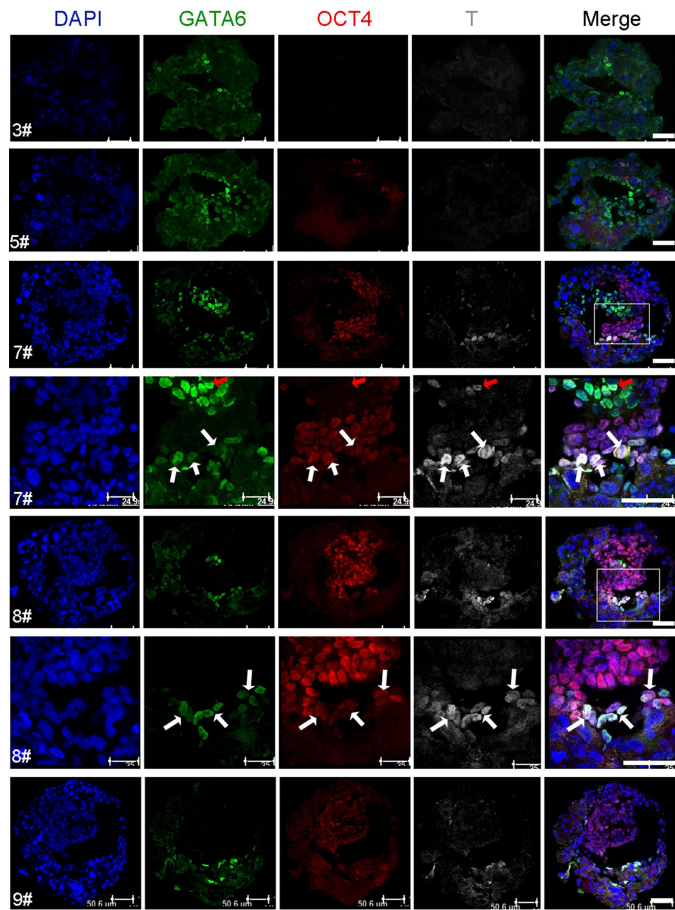
**l**, The violin plots show expression dynamics of primitive streak genes over pluripotent-stem-cell development. All violins have the same maximum width, black dot denotes the mean. In total, 136 cells were included (Extended Data Fig. 9e): ICM, *n* = 49 cells; pre-EPI, *n* = 23 cells; post-EPI, *n* = 48 cells; PSA-EPI, *n* = 16 cells. \**P* < 0.05, \*\**P* < 0.01, two-sided Wilcoxon rank-sum test. **m–o**, Absence of specific neural gene expression indicates 14-d.p.f. embryos do not generate the initial nervous system, which meets the internationally recognized ethical limit for human embryo culture. **m**, Violin plots of dynamic expressions of neural-specific genes in EPIs over embryo culture. All violins have the same maximum width, black dot denote the mean. In total, 136 cells were included: ICM, *n* = 49 cells; pre-EPI, *n* = 23 cells; post-EPI, *n* = 48 cells; PSA-EPI, *n* = 16 cells. **n**, **o**, Representative staining of PAX6, OCT4, SOX1 and FOXA2 in human 14-d.p.f. embryos (3/3 embryos). **p–r**, Development and cell proliferation of human embryos cultured in the Matr+10% Matr condition. Quantified data at each stage were based on five embryos from three independent experiments. Data are presented as mean ± s.d. **p**, Quantification of the dynamics of total cell number per embryo during culture. \**P* < 0.05, \*\**P* < 0.01, two-sided Student's *t*-test. **q**, Dynamics of OCT4<sup>+</sup> EPIs and GATA6<sup>+</sup> PrEs per embryo over culture. **r**, Dynamics of CK7<sup>+</sup> TrBs per embryo over culture. EPIs and PrEs maintained gradual proliferation at 8–10 d.p.f., after which their proliferation speeds accelerated. However, TrBs always maintained a rapid proliferation rate, which may be for establishing cell connections with the maternal environment. Scale bars, 25 μm.



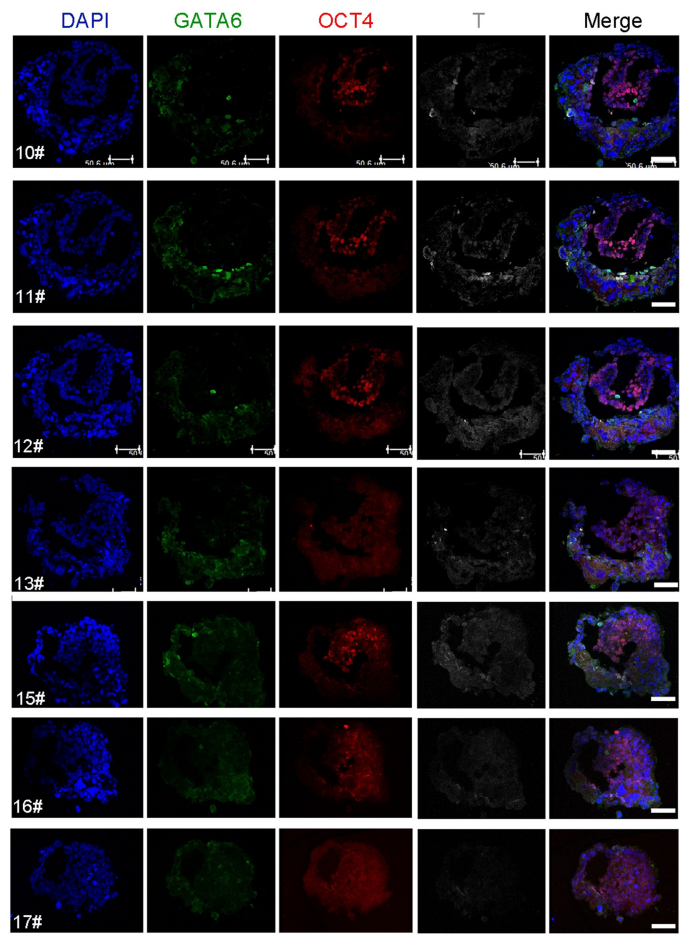


**Extended Data Fig. 6 | Z-series of a 14-d.p.f. embryo with a PSA.** This figure is related to Fig. 3. Series of confocal z-sections of the embryo stained for T (green), OCT4 (red) and N-cadherin (grey), showing formation of PSA. The thickness of every section was 12  $\mu$ m. Numbers on the left indicate the number of sections. Red and white arrows indicate T<sup>+</sup> and N-cadherin<sup>+</sup> cells, respectively. Some T<sup>+</sup> cells (red arrows) in the 13# section located in the amnion

epithelium, consistent with the distribution of some T<sup>+</sup> cells in the monkey amnion epithelium<sup>23</sup>. T<sup>+</sup> cells (red arrows) in the 16# section disrupted the N-cadherin-forming basement-membrane barriers between epiblast and hypoblast and focally migrated from the embryonic disc to generate the PSA. Similar data were observed in 3 out of 3 embryos from two experiments. Scale bars, 50  $\mu$ m.



**Extended Data Fig. 7 | Z-series of a 14-d.p.f. embryo with a cell emigration region.** This figure is related to Fig. 3. Series of confocal z-sections of the embryo stained for GATA6 (green), OCT4 (red) and T (grey). The thickness of every section was 12  $\mu\text{m}$ . Numbers on the left indicate the number of sections. Red and white arrows denote  $\text{T}^+\text{GATA6}^+\text{OCT4}^-$  cells and  $\text{T}^+\text{GATA6}^+\text{OCT4}^+$  cells,



respectively. In the 7# and 8# sections, some migrating  $\text{T}^+$  EPs invaded the space near the visceral endoderm and co-expressed GATA6 (white arrows). Similar data were observed in 4 out of 5 embryos from three experiments. Scale bars, 50  $\mu\text{m}$ .

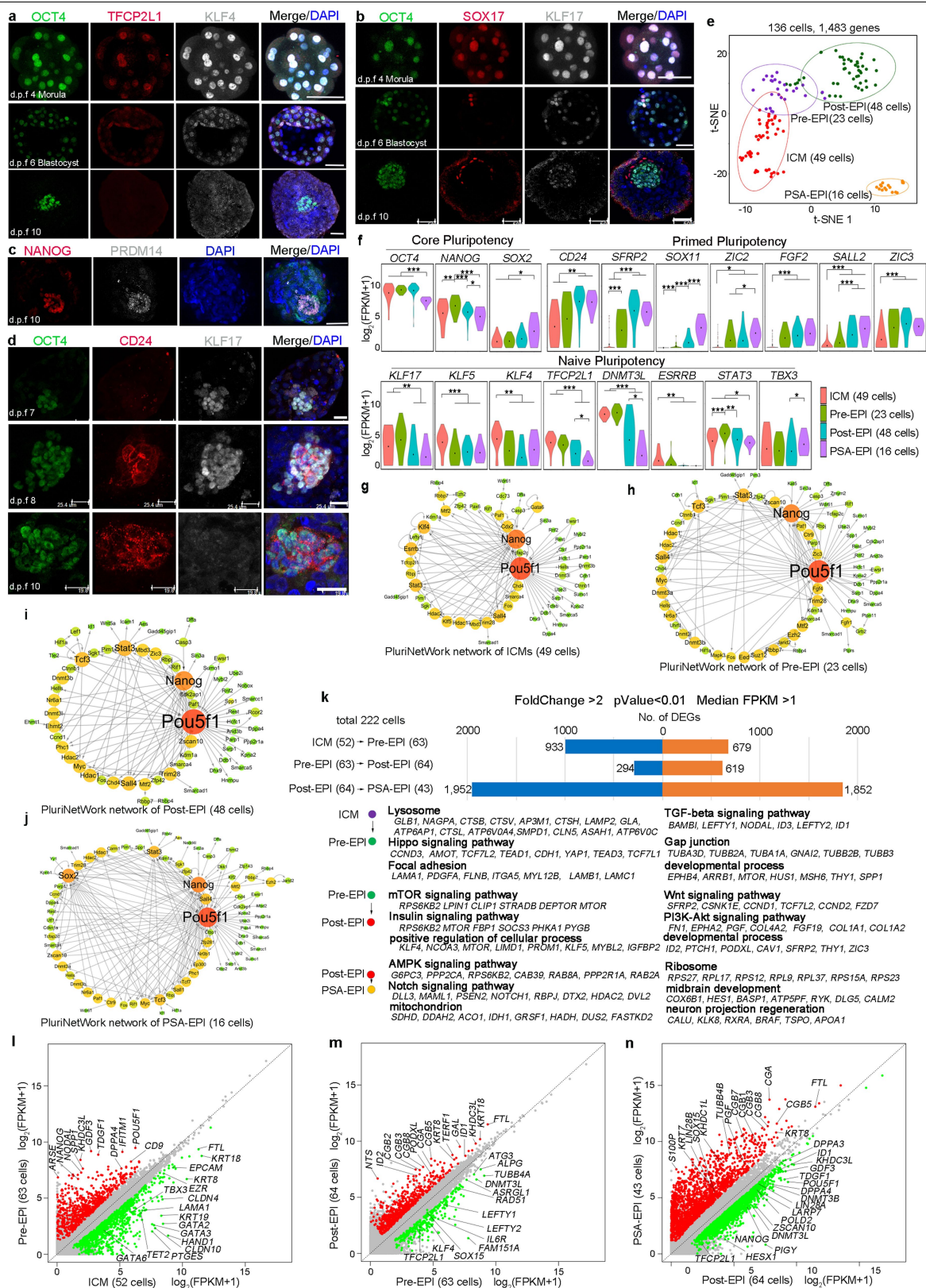




**Extended Data Fig. 8 | Development of TrB lineage. a–i,** Representative confocal staining of human embryo section. **a,** Staining of CK7 and F-actin.  $n = 5$  independent embryos from three independent experiments. **b, c,** Cells near the EPI–PrE bilayer had a single nucleus that expressed TEAD4 and E-cadherin. E-cadherin displayed symmetrical distributions in the embryos (3 out of 3 embryos for each staining from two experiments). **d, e,** Multinucleated cells (d1, e1 and e2) and cells with a single nucleus (d2) (3 out of 3 embryos each staining from two experiments). Dashed lines in **d** and **e** outline the region of interest in cells. **f, g,** Representative staining of CK7 and HLA-G (3 out of 3 embryos) in 12-d.p.f. (**f**) and 14-d.p.f. (**g**) human embryo. **h, i,** Representative staining of TEAD4 and HLA-G (**h**) or TEAD4 and hCG (**i**) in 14-d.p.f human embryo (3 out of 3 embryos). In **f**, the inset is magnification of the square. In **g–i**, right panel is magnification of the square. In **i**, the inset is magnification of the region indicated by a red arrow. **j–l, t-SNE** plot of 352 TrBs. Cells (dots) coloured according to the original inferred lineage identity (**k**) and embryonic stage (**l**). **m,** Lineage segregation path constructed by Monocle based on developmental

time (left) and cell types defined with selected markers (right). **n,** Heat map of polypeptide hormone genes expressed in the six types of trophoblasts during culture. **o,** Heat map indicates expression of genes specific for each cell type. Representative genes and key GO enrichments shown. GO terms and KEGG pathways of genes specific for different subtypes of TrBs (multiple test corrected  $P < 0.01$ , hypergeometric tests) in the six types of TrBs from pre-implantation stage embryos to 14-d.p.f embryos are shown in Supplementary Table 4. Pre-CTB-expressing genes related to cell metabolism, biosynthesis and cell differentiation, were in accord with the characteristics of trophoblast stem cells. High expressions of NF- $\kappa$ B, as well as canonical and non-canonical Wnt signalling pathway genes, indicate potential functions on CTB development. STB-specific genes indicated hormone secretion, whereas early-STB-expressing genes associated with cell differentiation and migration, dependent on several signalling pathways. EVT-specific genes contributed to regulating the immune system and angiogenesis. Scale bars, 50  $\mu$ m (**a–c, e–i**) or 20  $\mu$ m (**b**).





Extended Data Fig. 9 | See next page for caption.

**Extended Data Fig. 9 | Epiblast development during embryo culture in the 3D condition.** This figure relates to Fig. 4. **a–d**, Dynamic expressions of pluripotent genes over human embryo development in the 3D condition. **a**, The dynamic expressions of OCT4, TFCP2L1 and KLF4 during human embryo development (3 out of 3 embryos). **b**, The dynamic expressions of OCT4, SOX17 and KLF17 during human embryo development (3 out of 3 embryos). Loss of TFCP2L1, KLF4 and KLF17 at 10 d.p.f. (implantation stage) indicate the pluripotent state transition of epiblasts. **c**, Expression of NANOG and PRDM14 at 10-d.p.f. human embryos (2 out of 2 embryos). **d**, Dynamic expressions of OCT4, CD24 (a primed pluripotency gene) and KLF17 during human embryo development (2 out of 2 embryos). **e**, *t*-SNE of pluripotent stem cells. To exclude cells from the AME and intermediate state in the epiblast cluster, we excluded the NANOG-negative cells and maintained 136 cells with high expression of NANOG. **f**, The violin plots show the dynamics of naive, primed and common

pluripotency genes in pre-gastrulation embryos. All violins have the same maximum width, black dot denotes the mean. AME and intermediate state cells were excluded in the synthesis. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , two-sided Wilcoxon rank-sum test. **g–j**, PluriNetWork analysis of EPIs from ICM (**g**), pre-EPI (**h**), post-EPI (**i**) and PSA-EPI (**j**) stages revealed that key pluripotency regulators dominated the networks. **k**, DEGs during EPI development. GO terms and representative genes in DEGs in the pairwise comparisons are indicated. **l–n**, Scatter-plot comparison of the gene-expression levels between ICM and pre-EPI (**l**), pre-EPI and post-EPI (**m**), and post-EPI and PSA-EPI (**n**). Key genes are annotated (Supplementary Table 6). Red denotes upregulated; green denotes downregulated; >twofold difference, uncorrected  $P < 0.01$  (likelihood ratio test) and median FPKM > 1 in one group. Scale bars, 50  $\mu\text{m}$  (**a**, **b**), 100  $\mu\text{m}$  (**c**) or 25  $\mu\text{m}$  (**d**).



**Extended Data Fig. 10 | Comparison of monkey and human EPI development.**

**a**, PCA of the EPI lineage as determined by the expressed genes among all groups of EPIs during development in human (hu) (circles) and cynomolgus monkey (cy) (squares). In total, 12,475 out of 16,487 annotated gene expressed among human and monkey cells (human, 222 cells; monkey 213 cells) were used. **b**, Heat map of 966 genes that highly contributed to PC1 (>2 s.d. of PC1). **c**, Heat map of 1,152 genes with significant scores for PC2 and PC3 loading (radius of PC2 and PC3 > 3 s.d.) during monkey and human EPI development. None of the 1,151 genes overlapped with 996 genes with significant scores for PC1 loading. **d**, The violin plots of pluripotency genes over EPI pluripotency transition in monkey embryos. Monkey scRNA-seq data were from a published database<sup>39</sup>. **e**, The violin plots of pluripotency genes during the EPI pluripotency transition in human embryos. We observed stable expressions of

*STAT3* and *TBX3* with a trend of gradually increasing of *UTF1*, *NROB1*, *LIFR* and *SOX15* during human EPI pluripotent state transition. **f**, The violin plots showed dynamics of BMP signalling pathway gene expression in EPI pluripotency transition of human and monkey embryos. Monkey scRNA-seq data were obtained from a published database<sup>8</sup>. **g**, The violin plots showed dynamics of FGF signalling pathway gene expression in EPI pluripotency transition of human and monkey embryos. **h**, The violin plots showed dynamics of Notch signalling pathway gene expression in EPI pluripotency transition of human embryos. In **d-h**, all violins have the same maximum width, black dot denotes the mean. In **e-h**, AME and intermediate state cells were excluded in the synthesis and only included 136 single cells. Monkey scRNA-seq data were obtained from a published database<sup>8</sup>. See also Supplementary Table 7.



## Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- |                                     |  |
|-------------------------------------|--|
| n/a                                 | Confirmed  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided<br><i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A description of all covariates tested  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> For null hypothesis testing, the test statistic (e.g. $F$ , $t$ , $r$ ) with confidence intervals, effect sizes, degrees of freedom and $P$ value noted<br><i>Give <math>P</math> values as exact values whenever suitable.</i>                            |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings  |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> Estimates of effect sizes (e.g. Cohen's $d$ , Pearson's $r$ ), indicating how they were calculated   |

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

### Software and code

Policy information about [availability of computer code](#)

Data collection LAS X (2.6.0 build 7266) software were used to take pictures and 3D videos.

Data analysis The software used for scRNA-seq data analysis have been described in Materials and Methods, including the key parameter. Image J (version 1.51 j8) was used to count cell numbers; FASTQC (Version 0.11.8) and MULTIQC(v 1.6) were used for quality control; HISAT2(v 2.1.0) was used to align the scRNA-seq profiles to the human genome. Stringtie (v 1.3.4) was used to calculate the abundances of genes. Seurat (version 2.3.4) and Monocle (v 2.4.0) in R were used to perform t-SNE and trajectory analysis of the scRNA-seq profiles. For cell number, the significance difference between two samples was evaluated by unpaired two-sample Student's t-test using Excel software (2016).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

### Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Source data for Extended Data Fig.1f, k, 1l, 1m, 4b, 4f, 4p, 5g, 5h, 5l, 5m, 5p, 5q, 5r, 9f, 10d, 10e, 10f, 10g, 10h are provided with the paper. The single-cell RNA-sequencing data have been deposited in the GEO. Accession numbers for the data generated in this study and for the published data used in this study are as follows. The scRNA-seq data in this study: GSE136447; those SC3-seq data of cynomolgus monkey embryos (for Extended Data Fig.10), GSE74767 (ref.39); scRNA-seq data of human pre-implantation embryos (for Extended Data Fig.3d, e): GSE66507(ref.16), GSE36552 (ref.17) and E-MTAB-3929 (ref.18).

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical methods were used to predetermine sample size. The number of embryos used in each experiment were provided in methods and figure legends. For these experiments with some variations, we examined at least 3. For these experiments with highly consistent data, we used at least two embryos because of limited embryos. For RNA-Seq, the sample size was determined when the main cell lineages at each developmental stages were captured. Related statistical analysis provides the rationale for sufficiency of the sample sizes.
Data exclusions	Two scRNA-seq profiles whose numbers of genes with abundance levels more than 1 FPKM were smaller than 2000 were excluded. The criteria was established in the analysis procedure when we found that the two scRNA-seq profiles with < 2000 expressed genes were isolated far away from other scRNA-seq profiles.
Replication	Methods and figure legends indicated the exact number of embryos replicated in each experiment. All attempts at replication were successful.
Randomization	The embryos used in each experiment were chosen at random. The experiments were not randomized.
Blinding	The investigators were not blinded to allocation during experiments and outcome assessment. Data collection and analysis were performed by different people, the sample classification were replaced by simple marks during data analysis.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input type="checkbox"/>	<input checked="" type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data

### Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

## Antibodies

Antibodies used	Immunostaining: mouse anti-OCT3/4 (Santa Cruz, SC5279, C-10, H1612 1:400), rabbit anti-Brachyury (T) (Santa Cruz, SC20109, polyclonal, A0616, 1:50), rabbit anti-SOX2 (Millipore, AB5603, polyclonal, 2826070, 1:400), goat anti-SOX17 (R&D Systems, AF1924, polyclonal, KGA0815042, 1:250), rabbit anti-KLF4 (Millipore, 09-821, polyclonal, 2383578, 1:400), rabbit anti-KLF17 (Atlas Antibodies, HPA024629, polyclonal, C117502, 1:250), rabbit anti-β-Catenin (Abcam, AB32572, E247, GR184212-37, 1:300), mouse anti-E-cadherin (Abcam, AB76055, M168, GR299147-4, 1:100), rabbit anti-Cytokeratin 7 (CK7) (Abcam, AB181598, EPR17078, GR3214132-10, 1:300), rabbit anti-N-cadherin (Abcam, AB12221, polyclonal, 40975, 1:200), mouse anti-OTX2 (Santa Cruz, SC514195, D-8, G0816, 1:100), goat anti-LEFTY1 (R&D Systems, AF746, polyclonal, CMM0111101, 1:100), goat anti-SOX1 (R&D Systems, AF3369, polyclonal, XUV0618081, 1:500), rabbit anti-PAX6 (Biolegend, 901301, Poly 19013, b267205, 1:500), rabbit anti-Laminin (Sigma-Aldrich, L9393, polyclonal, 028M4890V, 1:50), rabbit anti-TEAD4 (Atlas Antibodies, HPA056896, polyclonal, R78063, 1:150), goat anti-FOXA2 (Santa Cruz, SC6554, polyclonal, D1216, 1:100), goat anti-OCT3/4 (Santa Cruz, SC8628, polyclonal, G3201, 1:250), Mouse anti-hCG (Abcam, AB9582, 5H4-E2, GR308272-2, 1:100), rabbit anti-PRDM14 (Millipore, AB4350, polyclonal, 2897240, 1:50), goat anti-TFCP2L1 (R&D Systems, AF5726, polyclonal, CCUG0115021, 1:200), goat anti-NANOG (R&D Systems, AF1997, polyclonal, KJ0514091, 1:250), mouse anti-PODXL (R&D Systems, MAB1658, 222328, JKW0218041, 1:400), mouse anti-HLA-G (Abcam, AB52455, 4H84, GR251679-19, 1:200), goat anti-GATA6 (R&D Systems, AF1700, polyclonal, KWT0419021, 1:200), mouse anti-EZRIN (Sigma-Aldrich, E8897, 3C12, 117M4875V, 1:500), rabbit anti-CER1 (Sigma-Aldrich, HPA019917, polyclonal, R10176, 1:50), Phalloidin (F-actin) Alexa Fluor®488 (Thermo Fisher Scientific, A12379, direct labeled, 1749905, 1:300), rabbit anti-HESX 1 (Abcam, AB246949, polyclonal, GR3267093-1, 1:100), goat anti-Brachyury (T)
-----------------	--

(R&D Systems, AF2085, polyclonal, KQP0617031, 1:200), goat anti FLK-1 (KDR) (R&D Systems, AF357, polyclonal, CVE0617081, 1:100) and WGA Alexa Fluor® 647 (Invitrogen, W32466, direct labeled, 1988457, 1:500)

## Validation

All the antibodies have been validated by the companies from which they were offered. This information was used for further validate of the antibodies used in this work. Details of the validation statements, antibody profiles and relevant citations can be found on the manufacturer's website.

All the antibodies used in this work are for immunostaining purpose only. We list the immunostaining validation of the manufactures and the number of citations as follows:

1. Oct3/4 (SC5279): <https://www.scbt.com/p/oct-3-4-antibody-c-10>  
Oct3/4 (SC5279) antibody was validated by the manufacturer using glandular cells and mouse embryos. We verified that this antibody stains the nuclei of ICM and EPIs, as expected. More than 1303 citations.
2. Brachyury(T)(SC20109): <https://www.scbt.com/zh/p/brachyury-antibody-h-210>  
Brachyury(T)(SC20109) antibody was validated by the manufacturer using human lung tissue. We verified that this antibody stains the nuclei of posterior EPIs, as expected. More than 9 citations.
3. Sox2 (AB5603): [http://www.merckmillipore.com/CN/zh/product/Anti-Sox2-Antibody,MM\\_NF-AB5603](http://www.merckmillipore.com/CN/zh/product/Anti-Sox2-Antibody,MM_NF-AB5603)  
Sox2 (AB5603) antibody was validated by the manufacture using H9 human stem cells. We verified that this antibody stains the nuclei of EPIs, as expected. More than 166 citations.
4. Sox17 (AF1924): [https://www.rndsystems.com/cn/products/human-sox17-antibody\\_af1924](https://www.rndsystems.com/cn/products/human-sox17-antibody_af1924)  
Sox17 (AF1924) antibody was validated by the manufacture using B16 mouse cell line and human BG01V cells. We verified that this antibody stains the nuclei of morula and PrEs, as expected. More than 123 citations.
5. Klf 4 (09-821): <http://www.merckmillipore.com/CN/zh/search/09-821>  
Klf 4 (09-821) antibody was validated by the manufacture's using NIH/3T3, A431, and HeLa cells. We verified that this antibody stains the nuclei of early EPIs. More than 4 citations.
6. Klf17(HPA024629):<https://www.atlasantibodies.com/products/antibodies/primary-antibodies/triple-a-polyclonals/klf17-antibody-hpa024629>  
Klf 17 (HPA024629) antibody was validated by the manufacture using human testis and tonsil tissues. We verified that this antibody stains the nuclei of early EPIs. More than 6 citations.
7.  $\beta$ -Catenin (AB32572): <https://www.abcam.cn/beta-catenin-antibody-e247-chip-grade-ab32572.html>  
 $\beta$ -Catenin (AB32572) antibody was validated by the manufacture using A431 and wild-type HAP1 cells. We verified that this antibody stains the membrane of the embryo cells. More than 335 citations.
8. E-cadherin (AB76055): <https://www.abcam.cn/e-cadherin-antibody-m168-c-terminal-ab76055.html>  
E-cadherin (AB76055) antibody was validated by the manufacture using A431 cells. We verified that this antibody stains the membrane of the embryo cells which N-cadherin did not expressed. More than 137 citations.
9. Cytokeratin7 (CK7) (AB181598) <https://www.abcam.cn/cytokeratin-7-antibody-epr17078-cytoskeleton-marker-ab181598.html>  
Cytokeratin7 (CK7) (AB181598) antibody was validated by the manufacture using A549 cells. We verified that this antibody stains the membrane of trophoblast. More than 24 citations.
10. N-cadherin (AB12221): <https://www.abcam.cn/n-cadherin-antibody-ab12221.html>  
N-cadherin (AB12221) antibody was validated by the manufacture using mouse differentiated embryonic stem cells. We verified that this antibody stains the membrane of mesenchymal cells. More than 84 citations.
11. Otx2 (SC514195): <https://www.scbt.com/zh/p/otx2-antibody-d-8>  
Otx2 (SC514195) antibody was validated by the manufacture using Jurkat, Hep G2, Hela nuclear extract and hES differentriated cells in our lab. We verified that this antibody stains the nuclei of parts of hypoblast cells. More than 2 citations.
12. Lefty1 (AF746): [https://www.rndsystems.com/cn/products/human-mouse-lefty-antibody\\_af746](https://www.rndsystems.com/cn/products/human-mouse-lefty-antibody_af746)  
Lefty1 (AF746) antibody was validated by the manufacture using mouse ulters in the citation. We verified that this antibody stains cytoplasma AVE cells, as expected. More than 1 citations.
13. Sox1 (AF3369): [https://www.rndsystems.com/cn/products/human-mouse-rat-sox1-antibody\\_af3369](https://www.rndsystems.com/cn/products/human-mouse-rat-sox1-antibody_af3369)  
Sox1 (AF3369) antibody was validated by the manufacture using ectoderm differentiated BG01V human embryonic stem cells. We verified that this antibody stains neural stem cells in our lab. More than 30 citations.
14. Pax6 (901301): <https://www.biolegend.com/en-us/products/purified-anti-pax-6-antibody-11511>  
Pax6 (901301) antibody was validated by the manufacture using frozen human iPSC derived neural rosettes. We verified that this antibody stains neural stem cells in our lab. More than 97 citations.
15. Laminin (L9393): <https://www.sigmaaldrich.com/catalog/product/sigma/l9393>  
Laminin (L9393) antibody was validated by the manufacture using human Tongue sections. We verified that this antibody stains basement membrane, as expected. More than 776 citations.
16. Tead4(HPA056896):<https://www.atlasantibodies.com/products/antibodies/primary-antibodies/triple-a-polyclonals/tead4-antibody-hpa056896/>  
Tead4(HPA056896) antibody was validated by the manufacture using human A431 cells. We verified that this antibody stains the nuclei of parts of EPIs and TEs, as expected. More than 4 citations.
17. Foxa2 (SC6554): <https://www.scbt.com/zh/p/hnf-3beta-antibody-m-20>  
Foxa2 (SC6554) antibody was validated by the manufacture using HepG2 cells. We verified that this antibody stains the nuclei of neural stem cells in our lab. More than 66 citations.
18. Oct3/4 (SC8628): <https://www.scbt.com/zh/p/oct-3-4-antibody-n-19>  
Oct3/4 (SC8628) antibody was validated by the manufacturer using mouse embryos. We verified that this antibody stains the nuclei of ICM and EPIs, as expected. More than 79 citations.
19. hCG (AB9582): <https://www.abcam.cn/hcg-beta-antibody-5h4-e2-ab9582.html>  
hCG (AB9582) antibody was validated by the manufacturer using human chorionic villus cells. We verified that this antibody stains the AME, as expected. More than 8 citations.
20. PRDM14 (AB4350): [http://www.merckmillipore.com/CN/zh/product/Anti-PRDM14-Antibody,MM\\_NF-AB4350#anchor\\_REF](http://www.merckmillipore.com/CN/zh/product/Anti-PRDM14-Antibody,MM_NF-AB4350#anchor_REF)  
PRDM14 (AB4350) antibody was validated by the manufacturer using human lung tissue lysate. We verified that this antibody stains the early EPIs, as expected. More than 1 citations.
21. TFPC2L1 (AF5726): [https://www.rndsystems.com/cn/products/human-tfcp2l1-antibody\\_af5726](https://www.rndsystems.com/cn/products/human-tfcp2l1-antibody_af5726)  
TFPC2L1 (AF5726) antibody was validated by the manufacturer using human placental tissue. We verified that this antibody stains the early EPIs and naïve stem cells in our lab. More than 5 citations.
22. NANOG (AF1997): [https://www.rndsystems.com/cn/products/human-nanog-antibody\\_af1997](https://www.rndsystems.com/cn/products/human-nanog-antibody_af1997)

NANOG (AF1997) antibody was validated by the manufacturer using BG01V human stem cells and embryoid body. We verified that this antibody stains the ICM and EPIs, as expected. More than 111 citations.

23. PODXL (MAB1658): [https://www.rndsystems.com/cn/products/human-podocalyxin-antibody-222328\\_mab1658](https://www.rndsystems.com/cn/products/human-podocalyxin-antibody-222328_mab1658)  
 PODXL (MAB1658) antibody was validated by the manufacturer using BG01V human stem cells. We verified that this antibody indicated the polarity of EPIs, as expected. More than 7 citations.

24. HLA-G (AB52455): <https://www.abcam.cn/hla-g-antibody-4h84-ab52455.html>  
 HLA-G (AB52455) antibody was validated by the manufacturer using human spleen tissue. We verified that this antibody stains the cytoplasm of parts of TEs, as expected. More than 11 citations.

25. GATA6 (AF1700): [https://www.rndsystems.com/cn/products/human-gata-6-antibody\\_af1700](https://www.rndsystems.com/cn/products/human-gata-6-antibody_af1700)  
 GATA6 (AF1700) antibody was validated by the manufacturer using human gastric carcinoma cell line. We verified that this antibody stains the nuclei of hypoblast, as expected. More than 29 citations.

26. EZRIN(E8897):<https://www.sigmaaldrich.com/catalog/product/sigma/e8897?lang=zh&region=CN>  
 EZRIN (E8897) antibody was validated by the manufacturer using A431 cells. We verified that this antibody expressed in apical surface in EPIs and AME, as expected. More than 40 citations.

27. CER1(HPA019917):<https://www.sigmaaldrich.com/catalog/product/sigma/hpa019917?lang=zh&region=CN>  
 CER1(HPA019917) antibody was validated by the manufacturer using human small intestine. We verified that this antibody expressed in AVE, as expected. More than 1 citations.

28. Phalloidin (F-actin) Alexa Fluor®488 (A12379): <https://www.thermofisher.com/order/catalog/product/A12379?SID=srch-hj-A12379>  
 Phalloidin (F-actin) Alexa Fluor®488 (A12379) antibody was validated by the manufacturer using bovine pulmonary artery endothelial cells and muntjac skin fibroblasts. We verified that this antibody expressed in AVE, as expected. More than 176 citations.

29. HESX 1 (AB246949): <https://www.abcam.cn/hesx1-antibody-ab246949.html>  
 HESX 1 (AB246949) antibody was validated by the manufacturer using human testis and skin tissue. We verified that this antibody stains the nuclei of part of EPIs, as expected. No citations at present.

30. Brachyury (T) (AF2085): [https://www.rndsystems.com/cn/products/human-mouse-brachyury-antibody\\_af2085](https://www.rndsystems.com/cn/products/human-mouse-brachyury-antibody_af2085)  
 Brachyury (T) (AF2085) antibody was validated by the manufacturer using differentiated human embryonic stem cells and BG01V human stem cells. We verified that this antibody stains the nuclei of posterior EPIs, as expected. More than 56 citations.

31. FLK-1 (KDR) (AF357): [https://www.rndsystems.com/cn/products/human-vegfr2-kdr-flk-1-antibody\\_af357](https://www.rndsystems.com/cn/products/human-vegfr2-kdr-flk-1-antibody_af357)  
 FLK-1 (KDR) (AF357) antibody was validated by the manufacturer using human placenta and kidney tissues. We verified that this antibody stains the extraembryonic mesoderm. More than 25 citations.

32. WGA Alexa Fluor® 647 (W32466): <https://www.thermofisher.com/order/catalog/product/W32466?SID=srch-hj-W32466>.  
 WGA Alexa Fluor® 647 (W32466) antibody was validated by citations using mouse cardiomyocytes on the manufacturer's website. We verified that this antibody expressed in apical surface in EPIs and AME. More than 4 citations.

## Human research participants

Policy information about [studies involving human research participants](#)

Population characteristics	All embryos collected are from the volunteers including wives and husbands who are no more than 35 years old. All volunteers have normal chromosome karyotype and no heredity case history. The embryo donors involved in this study are fertile with at least one healthy baby. The donors with normal body mass index are in good health. 86% couples were treated with normal IVF and 14% couples were treated with ICSI.
Recruitment	Research donors in the study were recruited from The First People's Hospital of Yunnan Province. Before giving consent, donors have received proper counselling about the implications of the donation and potential risks. Embryos were collected with written informed consent from the donors in this study.
Ethics oversight	This work was approved by the Medicine Ethics Committee of The First People's Hospital of Yunnan Province (2017LS[K]NO.035). The informed consent process for embryo donation complied with International Society for Stem Cell Research (ISSCR) Guidelines for Stem Cell Research and Clinical Translation (2016) and Ethical Guidelines for Human Embryonic Stem Cell research (2003) jointly issued by the Ministry of Science and Technology and the Ministry of Health of People's Republic of China. The Medicine Ethics Committee of The First People's Hospital of Yunnan Province is composed of 9 members, including lawyers, scientists and clinicians with relevant expertise. The Committee evaluated the scientific merit and ethical justification of this study and conducted a full review of the donations and use of these samples. No financial inducements were offered for the donations. In the process, couples were informed that their embryos would be used to study the developmental mechanisms of human embryos and that their donation would not affect their IVF cycle. The culture of all embryos was terminated at day 14 post-fertilization or upon the appearance of primitive streak anlage.

Note that full information on the approval of the study protocol must also be provided in the manuscript.



# Microbiota-targeted maternal antibodies protect neonates from enteric infection

<https://doi.org/10.1038/s41586-019-1898-4>

Received: 9 November 2018

Accepted: 1 November 2019

Published online: 8 January 2020

Wen Zheng<sup>1,5</sup>, Wenjing Zhao<sup>2,3,5</sup>, Meng Wu<sup>1</sup>, Xinyang Song<sup>1</sup>, Florence Caro<sup>3</sup>, Ximei Sun<sup>1</sup>, Francesca Gazzaniga<sup>1</sup>, Giuseppe Stefanetti<sup>1,4</sup>, Sungwhan Oh<sup>1</sup>, John J. Mekalanos<sup>3</sup> & Dennis L. Kasper<sup>1\*</sup>

Although maternal antibodies protect newborn babies from infection<sup>1,2</sup>, little is known about how protective antibodies are induced without prior pathogen exposure. Here we show that neonatal mice that lack the capacity to produce IgG are protected from infection with the enteric pathogen enterotoxigenic *Escherichia coli* by maternal natural IgG antibodies against the maternal microbiota when antibodies are delivered either across the placenta or through breast milk. By challenging pups that were fostered by either maternal antibody-sufficient or antibody-deficient dams, we found that IgG derived from breast milk was crucial for protection against mucosal disease induced by enterotoxigenic *E. coli*. IgG also provides protection against systemic infection by *E. coli*. Pups used the neonatal Fc receptor to transfer IgG from milk into serum. The maternal commensal microbiota can induce antibodies that recognize antigens expressed by enterotoxigenic *E. coli* and other Enterobacteriaceae species. Induction of maternal antibodies against a commensal *Pantoea* species confers protection against enterotoxigenic *E. coli* in pups. This role of the microbiota in eliciting protective antibodies to a specific neonatal pathogen represents an important host defence mechanism against infection in neonates.

Neonates are highly susceptible to microbial infections, not only because their immature immune system is less capable of generating adaptive immune effectors such as antibodies<sup>1,2</sup>, but also because they lack a diverse commensal microbiota that can antagonize pathogens independently of host responses<sup>3</sup>. Neonates acquire maternal antibodies through the placenta and through breast milk; however, in humans, antibodies derived from breast milk are dominated by secretory IgA antibodies, which are thought to exert their protective function on neonatal mucosal surfaces through mechanisms such as toxin or adhesin neutralization and bacterial agglutination<sup>4,5</sup>. Passive immunity to various pathogenic bacterial and viral infections (such as group B *Streptococcus*, *Haemophilus influenzae* and influenza viruses) can be transferred to neonates by maternal antigen-specific IgG antibodies induced by maternal colonization or vaccination<sup>6–8</sup>.

Although the benefits of maternal antibodies are widely accepted<sup>9</sup>, few studies have addressed whether maternal natural antibodies (mNabs)—that is, antibodies acquired without known exposure to the pathogen or through immunization—can help neonates to defend against pathogens. Although the commensal microbiota can shape the antibody repertoire<sup>10,11</sup>, how the diversity in mNabs is induced or how they mediate protection against infectious agents postnatally are unknown. Here we show that mNabs protect neonatal mice against both enteric and systemic infections with enterotoxigenic *E. coli* (ETEC). Notably, we found that the induction of mNabs depends on the commensal microbiota in pregnant dams. We show that a single commensal

species can induce cross-reactive mNabs that protects against ETEC in pups. In addition to acquisition through the placenta, pups can assimilate IgG mNabs directly from ingested milk into serum by a neonatal Fc receptor (FcRn)-dependent process. Our results provide insights into how the commensal microbiota of pregnant female mice drives antibody-dependent immunity in neonates through breast-feeding and demonstrate that protective IgG antibodies in breast milk act both locally and systemically.

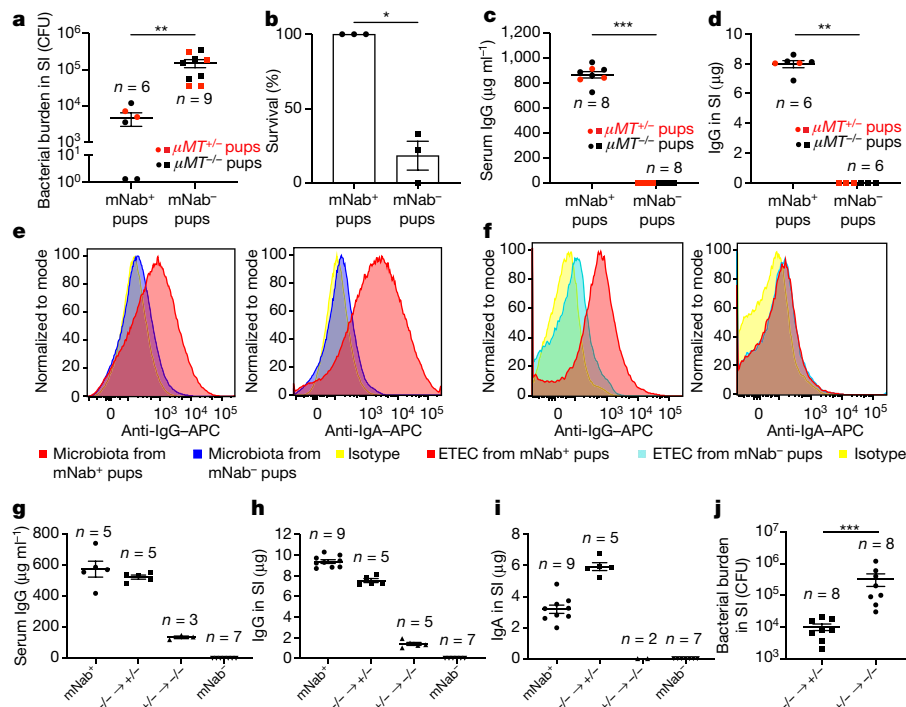
## Mouse mNabs protect neonates against ETEC

To analyse the developmental dynamics of neonatal antibodies, we used a reciprocal breeding strategy that enabled the tracking of maternal antibody persistence and antibody development dynamics in neonates. Maternal source, persistence and development of neonatal age-related IgG, IgA and IgM are shown in Extended Data Fig. 1. For the first 3 weeks, serum and mucosal IgG and IgA levels in pups depend completely on the maternal  $\mu$ MT (also known as *Ighm*) genotype ( $\mu$ MT<sup>−/−</sup> mice lack mature B cells). Through this breeding strategy, we can produce pups that are either deficient (mNab<sup>−</sup>) or sufficient (mNab<sup>+</sup>) in maternal natural IgG and IgA.

Transfer of vaccine-induced, antigen-specific antibodies confers passive protection in models of neonatal infection<sup>6,8,12</sup>. To test whether mNabs in unimmunized mice protect against an enteric pathogen, we challenged reciprocally bred 6- to 7-day-old pups with the human

<sup>1</sup>Department of Immunology, Harvard Medical School, Boston, MA, USA. <sup>2</sup>Guangdong Provincial Key Laboratory of Colorectal and Pelvic Floor Diseases, The Sixth Affiliated Hospital, School of Medicine, Sun Yat-sen University, Guangzhou, China. <sup>3</sup>Department of Microbiology, Harvard Medical School, Boston, MA, USA. <sup>4</sup>Department of Chemistry, University of Milan, Milan, Italy.

<sup>5</sup>These authors contributed equally: Wen Zheng, Wenjing Zhao. \*e-mail: dennis\_kasper@hms.harvard.edu



**Fig. 1 | mNabs protect neonates from an enteric bacterial pathogen.**

**a**, Bacterial burdens of reciprocally bred mNab<sup>+</sup> and mNab<sup>-</sup> pups (6–7 days old) orally challenged with  $10^7$  CFU of ETEC 6. Ig, immunoglobulin; SI, small intestine.  $^{**}P = 0.0004$ , two-tailed Mann–Whitney *U*-test. Data are representative of four independent experiments. **b**, Survival among reciprocally bred pups 20 h after oral–gastric challenge with  $10^9$  CFU of ETEC 6. Data are from three independent experiments (first experiment,  $n = 8$  mNab<sup>+</sup> mice,  $n = 5$  mNab<sup>-</sup> mice; second experiment,  $n = 9$  mNab<sup>+</sup> mice,  $n = 9$  mNab<sup>-</sup> mice; third experiment,  $n = 7$  mNab<sup>+</sup> mice,  $n = 6$  mNab<sup>-</sup> mice).  $^{*}P = 0.0011$ , two-tailed unpaired Student's *t*-test. **c**, Serum IgG levels in ETEC-challenged reciprocally bred pups.  $^{***}P = 0.0002$ , two-tailed Mann–Whitney *U*-test. Data are representative of two independent experiments. **d**, Small-intestinal mucosal IgG levels in ETEC-challenged reciprocally bred pups.  $^{**}P = 0.0022$ , two-tailed Mann–Whitney *U*-test. Data are representative of two independent

experiments. **e**, Flow cytometry analysis of natural maternal IgG and IgA coating of commensal bacteria of 1-week-old mNab<sup>+</sup> and mNab<sup>-</sup> pups. Data are representative of two independent experiments ( $n = 4$ – $5$  mice per group in each experiment). **f**, Flow cytometry analysis of natural maternal IgG and IgA coating of ETEC–GFP bacteria in mNab<sup>+</sup> and mNab<sup>-</sup> pups 18 h after infection. IgG and IgA signals are gated on GFP<sup>+</sup> population. Data are representative of two independent experiments ( $n = 4$ – $7$  mice per group in each experiment). **g**, Serum IgG levels after 1 week of cross-fostering. **h**, Small-intestinal IgG levels after 1 week of cross-fostering. **i**, Small-intestinal IgA levels after 1 week of cross-fostering. **j**, ETEC 6 bacterial burdens in the small intestine of pups cross-fostered for 1 week.  $^{***}P = 0.0002$ , two-tailed Mann–Whitney *U*-test. Data are representative of two independent experiments. **a–d, g–j**, Data are mean  $\pm$  s.e.m. Specific *n* numbers are indicated.

clinical isolate ETEC strain 6 (hereafter ETEC 6). ETEC 6 colonizes the small intestine of neonatal mice and typically causes acute and lethal diarrhoeal disease within 20 h of oral gastric challenge. At a sub-lethal dose of ETEC 6 ( $10^7$  colony-forming units; CFU), mNab<sup>+</sup> pups were more resistant to infection than mNab<sup>-</sup> pups and displayed a 33-fold reduction in intestinal colonization of ETEC 6 (Fig. 1a). Stratification by genotype showed no difference in bacterial burden between  $\mu MT^{+/-}$  and  $\mu MT^{-/-}$  pups. At a higher dose ( $10^9$  CFU), all mNab<sup>+</sup> pups were resistant to ETEC 6 challenge, whereas 83% of mNab<sup>-</sup> pups became moribund or had died within 20 h after challenge (Fig. 1b). The postnatal time of our ETEC challenge is too early for antigen-driven endogenous production of IgA and IgG; thus, the protective effects depend on maternally derived antibodies. We verified that IgG was detected in serum (Fig. 1c) and in gut luminal extracts (Fig. 1d) of only the mNab<sup>+</sup> pups. We also challenged reciprocally bred pups intraperitoneally and found that mNab<sup>+</sup> pups were more resistant to systemic infection with ETEC than mNab<sup>-</sup> pups (Extended Data Fig. 2a). Previous studies showed that natural IgM antibodies have broad specificity and provide protection against bacterial and viral infections<sup>13–17</sup>. However, natural IgM cannot be vertically transmitted from dams to pups (Extended Data Fig. 1e) and therefore is unlikely to play an important part in the protection against ETEC observed in our study.

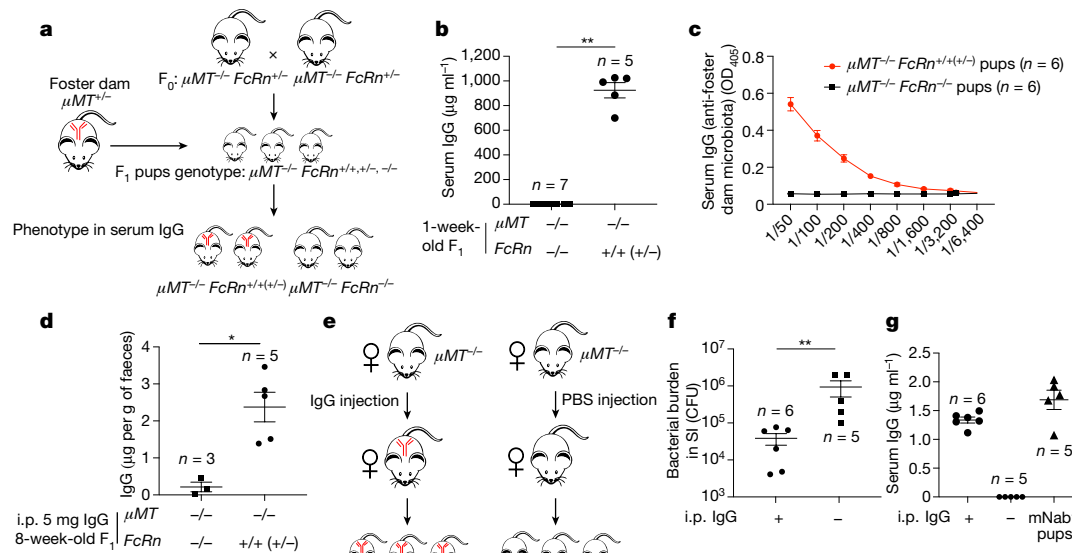
Using flow cytometry analysis, we investigated which antibody class was likely to mediate protection. Commensal bacteria from the

microbiota of uninfected mNab<sup>+</sup> pups were coated with both IgG and IgA, whereas bacterial cells from mNab<sup>-</sup> pups were negative for IgG and IgA (Fig. 1e), indicating that both maternal IgG and IgA that react with the commensal microbiota are transmitted vertically to neonates. Flow cytometry detected only IgG—but not IgA—on green-fluorescent protein (GFP)-expressing ETEC (ETEC–GFP) cells (Fig. 1f). It has previously been shown that immunization-induced antigen-specific milk IgG coats *Citrobacter* in the mucosa<sup>18</sup>. Our results re-affirm that maternal natural IgG in milk coats pathogenic bacterial cells (in our study ETEC 6), and further demonstrate that protection is conferred to breast-feeding pups, even without prior exposure of the mother to the pathogen.

The composition of the neonatal gut microbiota in mNab<sup>+</sup> and mNab<sup>-</sup> animals was similar and therefore probably not responsible for the differential protection against ETEC (Extended Data Fig. 2b). Exposure to maternal antibodies also suppressed the transcription of type-1 interferon-related genes in the small intestine of ETEC-infected pups (Extended Data Fig. 2c).

### Milk mNabs are critical to ETEC protection

To determine whether milk-acquired antibodies are protective, we orally challenged two groups of cross-fostered pups with  $10^9$  CFU of ETEC 6 (Extended Data Fig. 3a). Thus, mNab<sup>-</sup> pups were fostered by  $\mu MT^{+/-}$  dams and received their antibodies (designated  $\mu MT^{+/-}$ -to- $\mu MT^{+/-}$  pups) for



**Fig. 2 | FcRn mediates postnatal IgG retro-transport.** **a**, Breeding and fostering strategy to specifically study the postnatal milk IgG transfer process. All pups discussed in this figure are  $\mu MT^{-/-}$ . **b**, Serum IgG levels in 1-week-old FcRn-deficient or FcRn-sufficient  $\mu MT^{-/-}$  pups after 1 week of fostering by a  $\mu MT^{-/-}$  dam.  $**P=0.0013$ , two-tailed Mann–Whitney *U*-test. Data are representative of two independent experiments. **c**, Titres in pups of IgG specific to the microbiota of the foster dam. Data are representative of two independent experiments. **d**, Adult (8-week-old) mice were intraperitoneally (i.p.) injected with 5 mg of IgG, and faeces samples were collected 1 day later.

Faecal IgG levels are shown as microgram per gram of faeces.  $*P=0.0357$ , two-sided Mann–Whitney *U*-test. Data are representative of two independent experiments. **e**, IgG treatment scheme of dams. **f**, Comparison of ETEC 6 bacterial burden in the small intestine of pups from untreated  $\mu MT^{-/-}$  dams and from  $\mu MT^{-/-}$  dams treated with IgG.  $**P=0.0043$ , two-sided Mann–Whitney *U*-test. Data are representative of two independent experiments. **g**, Serum IgG levels of pups from untreated  $\mu MT^{-/-}$  dams compared with  $\mu MT^{-/-}$  dams treated with IgG. Data are representative of two independent experiments. **b–d, f, g**, Data are mean  $\pm$  s.e.m. Specific *n* numbers are indicated.

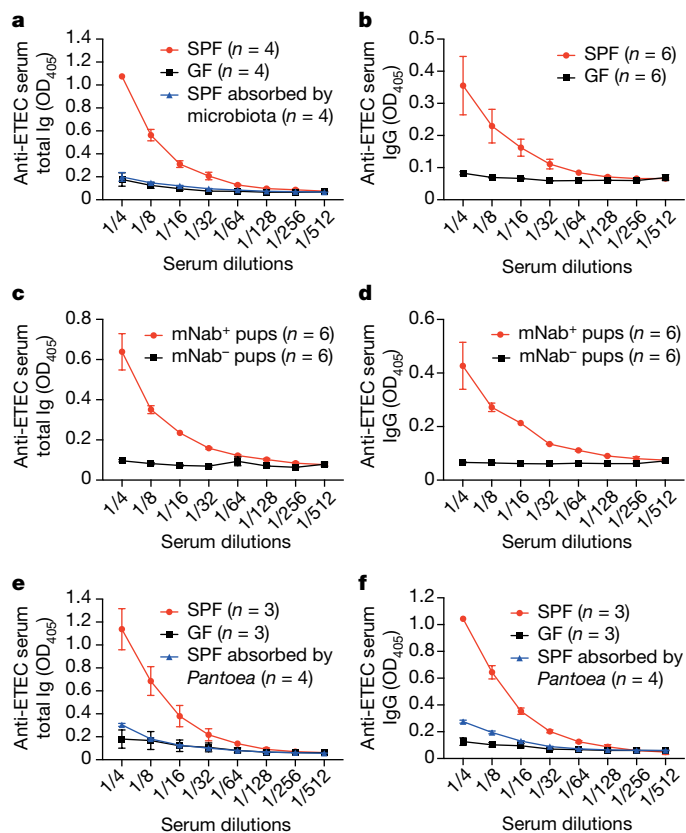
1 week only from milk, whereas mNab<sup>+</sup> pups were fostered by  $\mu MT^{-/-}$  dams (that is,  $\mu MT^{-/-}$ -to- $\mu MT^{-/-}$  pups). IgG titres of  $\mu MT^{-/-}$ -to- $\mu MT^{-/-}$  pups in the serum and small intestine were significantly higher than those of  $\mu MT^{-/-}$ -to- $\mu MT^{-/-}$  pups (Fig. 1g, h); the same trend was observed for IgA in the small intestine (Fig. 1i), and IgG and IgA in the colon (Extended Data Fig. 3b, c). After oral–gastric challenge of ETEC 6,  $\mu MT^{-/-}$ -to- $\mu MT^{-/-}$  pups had a bacterial burden in the small intestine that was approximately 30-fold lower than  $\mu MT^{-/-}$ -to- $\mu MT^{-/-}$  pups (Fig. 1j). In another cross-fostering experiment, pups born to  $\mu MT^{-/-}$  dams were divided into two groups and cross-fostered by  $\mu MT^{-/-}$  dams or their own  $\mu MT^{-/-}$  dams (Extended Data Fig. 3d). The mNab<sup>+</sup> pups fostered by  $\mu MT^{-/-}$  dams ( $\mu MT^{-/-}$ -to- $\mu MT^{-/-}$  pups) were resistant to ETEC 6 infection: 90% survived challenge at 20 h. By contrast, only 20% of mNab<sup>+</sup> pups raised by their own  $\mu MT^{-/-}$  dams survived (Extended Data Fig. 3e). This result confirms the importance of milk-derived mNabs in blocking ETEC 6 colonization and provides the mechanism by which milk-derived antibodies protect against ETEC 6 challenge.

## FcRn uptake of milk mNabs into the serum

In prenatal mice, FcRn transports IgG across the placenta to the fetus, primarily in the third trimester<sup>19</sup>. In adult mice, FcRn transports IgG from the gut lamina propria to the intestinal lumen; mice deficient in FcRn display impaired resistance to the enteric pathogen *Citrobacter rodentium*<sup>20</sup>. Because FcRn can have a high affinity for IgG and is expressed neonatally on intestinal epithelial cells<sup>21,22</sup>, we investigated whether FcRn transports IgG in the opposite direction, binding IgG in milk in the intestinal lumen and delivering it to the serum of suckling pups. We used a breeding and fostering strategy to separate the postnatal from prenatal antibody-transfer processes (Fig. 2a).  $\mu MT^{-/-} FcRn^{+/+}$  (also known as *Fcgrt*<sup>+/+</sup>) mice were used to generate littermate pups deficient in maternal antibodies. Newborn pups were immediately fostered by a  $\mu MT^{-/-}$  dam for 1 week.  $\mu MT^{-/-}$  pups sufficient in FcRn had almost 1 mg of IgG per ml of serum, whereas  $\mu MT^{-/-} FcRn^{-/-}$  pups had no

detectable serum IgG (Fig. 2b). In FcRn-sufficient pups, some IgG was directed towards the maternal microbiota (Fig. 2c). Thus, in suckling mice, FcRn transports IgG from ingested milk into the serum. Although FcRn can transport all subclasses of milk IgG to the neonatal circulation (Extended Data Fig. 4a–e), the relative serum concentrations of IgG3 and IgG2c (pup:dam ratios) are the highest and lowest, respectively (Extended Data Fig. 4f–j), suggesting that IgG3 is transferred preferentially and IgG2c the least efficiently. We then compared the role of FcRn in transferring IgG in adults versus neonates. We injected IgG (5 mg) intraperitoneally into 8-week-old littermates representing two groups:  $\mu MT^{-/-}$  (no antibody production) *FcRn*<sup>+/+</sup> or  $\mu MT^{-/-} FcRn^{-/-}$  (or *FcRn*<sup>+/+</sup>) mice. We sampled the faeces 1 day later.  $\mu MT^{-/-} FcRn^{+/+}$  (or  $\mu MT^{-/-} FcRn^{+/+}$ ) mice had significantly higher faecal IgG levels than  $\mu MT^{-/-} FcRn^{-/-}$  mice (Fig. 2d). Thus, IgG transfer from the systemic circulation to the intestinal lumen primarily depends on FcRn. However, when we orally gavaged IgG (5 mg) into these adult mice and sampled the serum 1 day later, we found that both  $\mu MT^{-/-} FcRn^{-/-}$  and  $\mu MT^{-/-} FcRn^{+/+}$  (or  $\mu MT^{-/-} FcRn^{+/+}$ ) mice had detectable but similarly low IgG levels. This experiment suggests that IgG transfer from lumen to serum in adult mice is poor and—in contrast to that in neonates—is not dependent on FcRn (Extended Data Fig. 4k); however, it is also possible that IgG given to adult mice by gavage was simply destroyed by proteolysis that does not occur in neonates.

IgG binding to the microbial surface can drive immune-effector functions such as complement-dependent bacteriolysis and opsonization, and because flow cytometry–analysed ETEC 6 cells were coated with IgG but not IgA in vivo, we hypothesized that maternal natural IgG was the immunoglobulin class that provided protection against ETEC. We synchronized the pregnancy of two  $\mu MT^{-/-}$  female mice mated with different  $\mu MT^{-/-}$  male mice. On gestational day 18 and postpartum day 2, one female received intraperitoneal injections of IgG (12 mg) purified from specific-pathogen-free (SPF) wild-type mice; the other received injections of only PBS (Fig. 2e). At 1 week of age, pups from these dams were challenged by the oral–gastric route with  $10^7$  CFU of ETEC 6. Pups



**Fig. 3 | The commensal microbiota elicits antibodies that cross-react with ETEC 6.** **a**, Total immunoglobulin titres against ETEC 6 in serum from germ-free (GF) and SPF adult female mice as well as in serum from SPF mice absorbed by mouse microbiota. Data are representative of four independent experiments. OD<sub>405</sub>, optical density at 405 nm. **b**, IgG titres against ETEC 6 in serum from germ-free and SPF mice. **c**, Total immunoglobulin titres against ETEC 6 in serum from 1-week-old neonatal mNab<sup>+</sup> and mNab<sup>-</sup> mice obtained by reciprocal breeding. **d**, IgG titres against ETEC 6 in serum from 1-week-old neonatal mNab<sup>+</sup> and mNab<sup>-</sup> mice obtained by reciprocal breeding. **e**, Total immunoglobulin titres against ETEC 6 in serum from germ-free mice, serum from SPF mice and *Pantoea*-1-absorbed serum from SPF mice. **f**, IgG titres against ETEC 6 in serum from germ-free mice, serum from SPF mice and *Pantoea*-1-absorbed serum from SPF mice. Data are mean  $\pm$  s.e.m. Specific *n* numbers are indicated in the figure.

born to the IgG-treated  $\mu$ MT<sup>-/-</sup> dam were highly protected and carried an approximately 25-fold lower small-intestinal bacterial load than pups born to the untreated dam ( $3.8 \times 10^4$  CFU per small intestine compared with  $9.4 \times 10^5$  CFU per small intestine; Fig. 2f). Thus IgG, in the absence of IgA, provides measurable protection against ETEC 6 challenge in nursing pups. We measured serum IgG titres in  $\mu$ MT<sup>-/-</sup> pups fed on a  $\mu$ MT<sup>-/-</sup> dam given the same passive IgG treatment and found titres that were comparable to those found in pups born to a  $\mu$ MT<sup>-/-</sup> dam (Fig. 2g). Thus, supplementing IgG antibodies to a pregnant or postpartum dam is sufficient to protect the pups that she nurses from ETEC 6 infection.

### Commensal *Pantoea* elicits mNabs protective against ETEC

The protection against ETEC of pups born to or nursed by mNab<sup>+</sup> dams suggests that conventionally colonized (SPF) mice carry cross-reacting natural antibodies against ETEC 6. Indeed, total ETEC-6-directed serum immunoglobulin and IgG titres are significantly higher in SPF mice than in germ-free mice (Fig. 3a, b), suggesting that the SPF commensal microbiota induces antibodies that cross-react with ETEC 6 in dams. Absorption of SPF mouse sera with faecal bacteria completely removed ETEC cross-reactive antibodies (Fig. 3a). At 1 week of age, mNab<sup>+</sup> pups

(born to  $\mu$ MT<sup>-/-</sup> dam) had substantial serum titres of ETEC-6-specific total immunoglobulin and IgG antibodies—presumably cross-reacting antibodies generated to commensal antigens—whereas sera from mNab<sup>-</sup> ( $\mu$ MT<sup>-/-</sup> dam) pups contained no detectable immunoglobulin that bound to this strain (Fig. 3c, d). Gram-negative bacteria of the family Enterobacteriaceae were isolated through the culture of faeces from  $\mu$ MT<sup>-/-</sup> dams. Mice from both our in-house breeding facility and The Jackson Laboratory lacked viable lactose-fermenting Gram-negative bacteria such as *E. coli* in their faeces. We could isolate only two lactose-non-fermenting Gram-negative Enterobacteriaceae species—a *Pantoea* and an *Enterobacter* species, each identified by 16S rRNA gene sequencing—from  $\mu$ MT<sup>-/-</sup> dams. We used the *Pantoea* 1 strain to absorb antibodies from mouse serum. *Pantoea*-1-absorbed SPF serum showed reduced titres of ETEC-reactive total immunoglobulin and IgG (Fig. 3e, f). These data support the hypothesis that some commensal microbiota species elicit cross-reactive antibodies against ETEC.

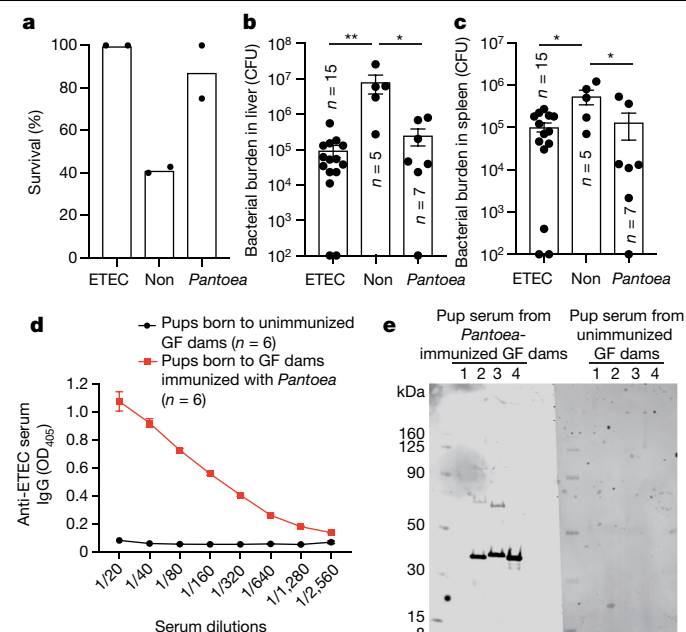
We wondered whether mNabs in mice showed cross-reactivity with other common enteric bacteria, including other pathogens and probiotic microorganisms. We measured the titres of antibodies in sera of germ-free and SPF mice that recognize *E. coli* Nissle, a human commensal bacterial isolate that has been used as a probiotic and is not present in the mouse gut, or *Salmonella typhimurium*, a pathogen in both humans and mice. Sera from SPF mice have higher total immunoglobulin titres to *E. coli* Nissle and *S. typhimurium* than sera of germ-free mice (Extended Data Fig. 5a, b). This result suggests that some commensal strains of the phylum Proteobacteria induce antibodies that recognize other proteobacterial species and strains.

To determine whether a single commensal species is sufficient to confer protection against ETEC 6 infection, we immunized germ-free dams either with a formalin-killed commensal *Pantoea* 1 strain or with a formalin-killed ETEC 6 strain and used unimmunized mice as a control group; then all three groups of pups were infected with ETEC 6. We found that pups born to germ-free dams immunized with *Pantoea* 1 were significantly more protected against ETEC 6 than pups born to unimmunized germ-free dams (Fig. 4a–c). IgG collected from pups born to *Pantoea*-1-immunized germ-free dams showed cross-reactivity to ETEC 6 (Fig. 4d) and the enteric pathogen *C. rodentium* (Extended Data Fig. 6). Furthermore, all of the commensal Enterobacteriaceae family isolates from mice found in three different vivariums were cross-reactive with the *Pantoea* anti-serum, but did not react with mouse or human commensal strains of *Staphylococcus* or *Bacteroides*. Pups of germ-free unimmunized dams had no detectable antibodies against ETEC 6 or these other bacterial species (Fig. 4e and Extended Data Fig. 6a, b). Western blot analysis of pronase-treated bacterial lysates showed elimination of a band cross-reactive with anti-*Pantoea* IgG, suggesting that this immunoreactive material is a protein (Extended Data Fig. 6c). We also measured IgG and IgA antibody content in milk samples from conventional SPF mice and found an IgG concentration that was approximately threefold higher than the concentration of IgA (Extended Data Fig. 6d); IgG titres in the milk of a given mouse dam were higher against the stool microbiota of the homologous dam than against that of a heterologous dam (Extended Data Fig. 6e). Collectively, these data suggest that the commensal microbiota can induce cross-reactive, protective antibodies against pathogens.

### Discussion

Of the many causes of death due to bacterial pathogens among children under 5 years old, acute infectious diarrhoea is surpassed only by pneumonia<sup>23</sup>. Neonates in developing countries have frequent diarrhoeal episodes that result in high mortality rates; the major infectious agents, which account for around 1.5 million deaths annually, are ETEC, rotavirus, *Vibrio cholerae* and *Shigella*<sup>24,25</sup>. ETEC is a frequent cause of diarrhoea in infants under 2 years old<sup>26</sup>. Epidemiological data indicate that breast-feeding reduces overall rates of diarrhoea and mortality<sup>26,27</sup>;





**Fig. 4 | Immunization of dams with commensal microorganisms conveys neonatal protection against pathogens.** **a**, Survival of pups born to ETEC-6- or *Pantoea*-1-immunized dams or unimmunized dams. Data are from two individual experiments (first experiment, ETEC  $n = 12$  mice, unimmunized (non)  $n = 7$  mice, *Pantoea*  $n = 4$  mice; second experiment, ETEC  $n = 5$  mice, unimmunized  $n = 5$  mice, *Pantoea*  $n = 4$  mice). **b**, Liver total bacterial burdens 3 days after intraperitoneal ETEC 6 challenge.  $**P = 0.0025$ , one-way analysis of variance (ANOVA) with Bonferroni post-test. Data are from two independent experiments. **c**, Spleen bacterial burdens 3 days after intraperitoneal ETEC 6 challenge.  $**P = 0.0041$ , one-way ANOVA with Bonferroni post-test. Data are from two independent experiments. **d**, Cross-reactivity against ETEC 6 of serum IgG from pups born to germ-free dams with or without *Pantoea* 1 immunization. **e**, Western blot showing that serum IgG of pups born to a *Pantoea*-1-immunized dam recognizes antigens in cellular lysates of members of the Enterobacteriaceae family (ETEC 6, *Pantoea* 1 and *Enterobacter*). Lane 1, *Staphylococcus*; lane 2, ETEC 6; lane 3, *Pantoea* 1; lane 4, *Enterobacter*. Blot is detected with goat anti-mouse IgG antibody. Data are representative of three independent experiments. For gel source data, see Supplementary Fig. 1. **b–d**, Data are mean  $\pm$  s.e.m.

however, the underlying mechanisms by which breast milk provides protection are not clear. Our results suggest that breast-feeding by mothers who lack specific immunity to ETEC may protect infants from ETEC by delivering natural antibodies—which are elicited by the commensal microbiota—that cross-react with this pathogen. The data presented here on cross-species protection by antibodies generated to a commensal organism are all based on mouse studies; further studies to address their relevance in humans are important.

IgG present in the breast milk of a selected dam reacted more strongly with the microbiota of that dam than with microbiota of other dams. We hypothesize that commensal species probably vary in their ability to induce cross-reacting antibodies that recognize any given pathogen, such as ETEC. Beyond antigens shared by specific bacterial groups (for example, lipopolysaccharides of Gram-negative bacteria), some antigens can be expressed by diverse and phylogenetically distant bacterial species, including commensal microorganisms<sup>28</sup>. Moreover, poly-reactive IgM can recognize both pathogenic and commensal bacteria and affords some protection against pathogen challenge in mice<sup>29</sup>. Thus, our study indicates that modulation of the maternal microbiota to optimize the induction of cross-reactive antibodies that are protective against important neonatal pathogens should be explored.

The role of secretory IgA in humoral responses to enteric pathogens has been widely studied<sup>30</sup>. The function of other antibody classes

(for example, IgG) at mucosal sites or in breast milk has attracted less attention, primarily because IgG is thought to be present at lower concentrations and to be less stable in mucosal secretions<sup>31–33</sup>. The consensus has been that secretory IgA in breast milk probably mediates protection<sup>34</sup>. Human and rodent milk contains substantial amounts of both secretory IgA and IgG<sup>35,36</sup>. In mice, microbiota-induced maternal IgG in milk is present in the neonatal gut mucosa and is taken up into the serum of breast-feeding neonates. Thus, breast-feeding theoretically could provide IgG-mediated protection against invasive pathogenic bacterial species at sites at which the effector mechanisms function, such as mucosal or submucosal surfaces, bloodstream or deeper tissues.

One important, previously unresolved question was whether orally delivered IgG (acquired by neonates through the milk of the mother) enters the bloodstream through a specific IgG transporter. Previous studies detected such transport for certain immunoglobulin classes<sup>10</sup> but did not clearly define the pathway for uptake. In vitro studies yielded evidence that FcRn recognizes IgG and transports it bidirectionally across an epithelial monolayer<sup>37</sup>. Our work in mice suggests that, dependent on FcRn, IgG in milk can enter the bloodstream of neonatal mice and confer potent protection—presumably through IgG-dependent effector functions such as complement classical pathway-dependent bacteriolysis or opsonization<sup>38</sup>. We also uncovered an FcRn-dependent pathway for retro-transport of IgG (Extended Data Fig. 7) relative to secretory processes mediated by the polymeric immunoglobulin receptor, which transports IgA and IgM from the basolateral to the apical surface and lumen of the intestine<sup>39</sup>. In the MDCK cell line, luminal-to-basolateral IgG transport reportedly requires antibody–antigen complexes and FcRn<sup>40</sup>; we did not observe transport of IgG from lumen to serum in adult mice. Although FcRn is thought to function bidirectionally, we observed that, in the mouse, the bidirectionality may be subject to an age-dependent temporal sequence (that is, in neonates, from lumen to submucosa; in adults, from submucosa to lumen). Characterizing this transport pathway in humans is a future priority because vaccination of women may generate high-affinity IgG, protecting breast-fed neonates long after antibodies received through the placenta have waned from the bloodstream. We have not addressed whether this milk-mediated gastrointestinal pathway for introducing therapeutic or preventive IgG into the bloodstream is applicable to human neonatal infants. If efficient and practical, this non-invasive approach offers advantages over conventional passive-immunization strategies by avoiding needle use in newborns, a practice that carries additional risk of disease transmission.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-019-1898-4>.

1. Basha, S., Surendran, N. & Pichichero, M. Immune responses in neonates. *Expert Rev. Clin. Immunol.* **10**, 1171–1184 (2014).
2. Simon, A. K., Hollander, G. A. & McMichael, A. Evolution of the immune system in humans from infancy to old age. *Proc. R. Soc. B* **282**, 20143085 (2015).
3. Kamada, N., Chen, G. Y., Inohara, N. & Núñez, G. Control of pathogens and pathobionts by the gut microbiota. *Nat. Immunol.* **14**, 685–690 (2013).
4. Carbonare, C. B., Carbonare, S. B. & Carneiro-Sampaio, M. M. S. Secretory immunoglobulin A obtained from pooled human colostrum and milk for oral passive immunization. *Pediatr. Allergy Immunol.* **16**, 574–581 (2005).
5. Hanson, L. A. R. & Korotkova, M. The role of breastfeeding in prevention of neonatal infection. *Semin. Neonatol.* **7**, 275–281 (2002).
6. Madoff, L. C., Michel, J. L., Gong, E. W., Rodewald, A. K. & Kasper, D. L. Protection of neonatal mice from group B streptococcal infection by maternal immunization with beta C protein. *Infect. Immun.* **60**, 4989–4994 (1992).
7. Zaman, K. et al. Effectiveness of maternal influenza immunization in mothers and infants. *N. Engl. J. Med.* **359**, 1555–1564 (2008).

8. Englund, J. A. et al. Transplacental antibody transfer following maternal immunization with polysaccharide and conjugate *Haemophilus influenzae* type b vaccines. *J. Infect. Dis.* **171**, 99–105 (1995).
9. Kearney, J. F., Patel, P., Stefanov, E. K. & King, R. G. Natural antibody repertoires: development and functional role in inhibiting allergic airway disease. *Annu. Rev. Immunol.* **33**, 475–504 (2015).
10. Macpherson, A. J., de Agüero, M. G. & Ganai-Vonarburg, S. C. How nutrition and the maternal microbiota shape the neonatal immune system. *Nat. Rev. Immunol.* **17**, 508–517 (2017).
11. Chen, Y. et al. Microbial symbionts regulate the primary Ig repertoire. *J. Exp. Med.* **215**, 1397–1415 (2018).
12. Englund, J. A. et al. Maternal immunization with influenza or tetanus toxoid vaccine for passive antibody protection in young infants. *J. Infect. Dis.* **168**, 647–656 (1993).
13. Boes, M., Prodeus, A. P., Schmidt, T., Carroll, M. C. & Chen, J. A critical role of natural immunoglobulin M in immediate defense against systemic bacterial infection. *J. Exp. Med.* **188**, 2381–2386 (1998).
14. Ochsenbein, A. F. et al. Control of early viral and bacterial distribution and disease by natural antibodies. *Science* **286**, 2156–2159 (1999).
15. Baumgarth, N. et al. B-1 and B-2 cell-derived immunoglobulin M antibodies are nonredundant components of the protective response to influenza virus infection. *J. Exp. Med.* **192**, 271–280 (2000).
16. Jayasekera, J. P., Moseman, E. A. & Carroll, M. C. Natural antibody and complement mediate neutralization of influenza virus in the absence of prior immunity. *J. Virol.* **81**, 3487–3494 (2007).
17. Zhou, Z. H. et al. The broad antibacterial activity of the natural antibody repertoire is due to polyreactive antibodies. *Cell Host Microbe* **1**, 51–61 (2007).
18. Caballero-Flores, G. et al. Maternal immunization confers protection to the offspring against an attaching and effacing pathogen through delivery of IgG in breast milk. *Cell Host Microbe* **25**, 313–323 (2019).
19. Palmeira, P., Quinello, C., Silveira-Lessa, A. L., Zago, C. A. & Carneiro-Sampaio, M. IgG placental transfer in healthy and pathological pregnancies. *Clin. Dev. Immunol.* **2012**, 985646 (2012).
20. Masuda, A. et al. Fcγ receptor regulation of *Citrobacter rodentium* infection. *Infect. Immun.* **76**, 1728–1737 (2008).
21. Pyzik, M., Rath, T., Lencer, W. I., Baker, K. & Blumberg, R. S. FcRn: the architect behind the immune and nonimmune functions of IgG and albumin. *J. Immunol.* **194**, 4595–4603 (2015).
22. Israel, E. J. et al. Expression of the neonatal Fc receptor, FcRn, on human intestinal epithelial cells. *Immunology* **92**, 69–74 (1997).
23. Kotloff, K. L. et al. Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): a prospective, case–control study. *Lancet* **382**, 209–222 (2013).
24. Kotloff, K. L. et al. Global burden of diarrheal diseases among children in developing countries: incidence, etiology, and insights from new molecular diagnostic techniques. *Vaccine* **35**, 6783–6789 (2017).
25. Kotloff, K. L. et al. The incidence, aetiology, and adverse clinical consequences of less severe diarrhoeal episodes among infants and children residing in low-income and middle-income countries: a 12-month case–control study as a follow-on to the Global Enteric Multicenter Study (GEMS). *Lancet Glob. Health* **7**, e568–e584 (2019).
26. Qadri, F., Svennerholm, A.-M., Faruque, A. S. G. & Sack, R. B. Enterotoxigenic *Escherichia coli* in developing countries: epidemiology, microbiology, clinical features, treatment, and prevention. *Clin. Microbiol. Rev.* **18**, 465–483 (2005).
27. Thapar, N. & Sanderson, I. R. Diarrhoea in children: an interface between developing and developed countries. *Lancet* **363**, 641–653 (2004).
28. Skurnik, D., Cywes-Bentley, C. & Pier, G. B. The exceptionally broad-based potential of active and passive vaccination targeting the conserved microbial surface polysaccharide PNAG. *Expert Rev. Vaccines* **15**, 1041–1053 (2016).
29. Le Gallou, S. et al. A splenic IgM memory subset with antibacterial specificities is sustained from persistent mucosal responses. *J. Exp. Med.* **215**, 2035–2053 (2018).
30. Wilmore, J. R. et al. Commensal microbes induce serum IgA responses that protect against polymicrobial sepsis. *Cell Host Microbe* **23**, 302–311 (2018).
31. Apter, F. M. et al. Analysis of the roles of antilipopolysaccharide and anti-cholera toxin immunoglobulin A (IgA) antibodies in protection against *Vibrio cholerae* and cholera toxin by use of monoclonal IgA antibodies in vivo. *Infect. Immun.* **61**, 5279–5285 (1993).
32. Michetti, P., Mahan, M. J., Schlauch, J. M., Mekalanos, J. J. & Neutra, M. R. Monoclonal secretory immunoglobulin A protects mice against oral challenge with the invasive pathogen *Salmonella typhimurium*. *Infect. Immun.* **60**, 1786–1792 (1992).
33. Moor, K. et al. High-avidity IgA protects the intestine by enchainning growing bacteria. *Nature* **544**, 498–502 (2017).
34. Stuebe, A. The risks of not breastfeeding for mothers and infants. *Rev. Obstet. Gynecol.* **2**, 222–231 (2009).
35. Goldsmith, S. J., Dickson, J. S., Barnhart, H. M., Toledo, R. T. & Eiten-Miller, R. R. IgA, IgG, IgM and lactoferrin contents of human milk during early lactation and the effect of processing and storage. *J. Food Prot.* **46**, 4–7 (1983).
36. Fouda, G. G. et al. HIV-specific functional antibody responses in breast milk mirror those in plasma and are primarily mediated by IgG antibodies. *J. Virol.* **85**, 9555–9567 (2011).
37. Dickinson, B. L. et al. Bidirectional FcRn-dependent IgG transport in a polarized human intestinal epithelial cell line. *J. Clin. Invest.* **104**, 903–911 (1999).
38. Bournazos, S. & Ravetch, J. V. Diversification of IgG effector functions. *Int. Immunol.* **29**, 303–310 (2017).
39. Mostov, K. E. Trans epithelial transport of immunoglobulins. *Annu. Rev. Immunol.* **12**, 63–84 (1994).
40. Yoshida, M. et al. Human neonatal Fc receptor mediates transport of IgG into luminal secretions for delivery of antigens to mucosal dendritic cells. *Immunity* **20**, 769–783 (2004).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2020

## Methods

### Data reporting

No statistical methods were used to predetermine sample size. The investigators were blinded to allocation during most of the experiments and outcome assessments.

### Mouse breeding strategy

Reciprocal breeding was used to generate pups that were sufficient or deficient in mNabs.  $\mu MT^{-/-}$  pups in the mNab<sup>+</sup> group were used to evaluate the persistence of mNabs and  $\mu MT^{-/-}$  pups in the mNab<sup>-</sup> group were used to evaluate the emergence of endogenous antibodies.

$\mu MT^{-/-}$  mice (stock no. 002288) and wild-type C57BL/6J mice (stock no. 000664) were purchased from The Jackson Laboratory and bred to generate  $F_1 \mu MT^{+/-}$  mice.  $F_1 \mu MT^{+/-}$  female mice were bred with  $\mu MT^{+/-}$  males to generate  $F_2$  progeny.  $F_2$  or  $F_3 \mu MT^{-/-}$  female  $\times \mu MT^{-/-}$  male breeding and  $\mu MT^{-/-}$  female  $\times \mu MT^{+/-}$  male breeding were synchronized to generate mNab<sup>+</sup> pups as well as mNab<sup>-</sup> pups. These pups were used for studies of ETEC 6 infection, in which serum and mucosal antibody levels were measured.  $FcRn^{-/-}$  mice (Jackson Laboratory, stock no. 003982) and  $\mu MT^{-/-}$  mice were used to generate the  $F_1 \mu MT^{+/-} FcRn^{+/-}$  progeny.  $F_1 \mu MT^{+/-} FcRn^{+/-}$  mice were then used to generate  $F_2 \mu MT^{-/-} FcRn^{+/-}$  progeny.  $F_2$  mice were used to generate  $\mu MT^{-/-} FcRn^{-/-}$  or  $\mu MT^{-/-} FcRn^{+/-}$  (or  $\mu MT^{-/-} FcRn^{+/+}$ ) mice. Germ-free C57BL/6J mice were bred and maintained in mouse facilities at Harvard Medical School. Germ-free mice were housed in standard isolators and were free of all bacteria, fungi, viruses and parasites; sterility was verified by regular interval aerobic and anaerobic cultures as well as PCR. All animal studies were approved by the IACUC of Harvard Medical School under animal protocol IS:00000178-3. Mouse genotyping followed the Jackson Laboratory genotyping protocol for stock no. 002288 and for stock no. 003982 (The Jackson Laboratory).

### Microbiota composition analysis

Faecal contents were scraped off the intestines of 1-week-old pups and DNA was extracted with a QIAamp DNA Stool Mini Kit (Qiagen, 51604). The V4 region of 16S rRNA gene was amplified with paired-end 16S rRNA gene primers 515F and 806R<sup>41</sup>, and approximately 390-bp amplicons were purified and then subjected to multiplex sequencing (Illumina MiSeq, 251 nucleotides  $\times$  2 paired-end reads with 12-nucleotide index reads). Raw sequencing data were analysed with QIIME2 pipelines<sup>42</sup>. The feature table of gut microbiota was then used for alpha and beta diversity analysis, as well as taxonomic analysis and differential abundance testing.

### Enteric pathogen infection of neonatal mice

In the intestinal infection model, to estimate bacterial burden, *E. coli* strain ETEC 6 ( $10^7$  CFU in 50  $\mu$ l PBS buffer) was administered orally by gavage to 6-day-old pups using an insulin needle connected to polyethylene tubing (Intramedic, 4274010). The ETEC 6 strain was a gift from F. Qadri; genome sequence, NCBI BioSample Accession number SAMN12263012.) Animals were monitored closely and euthanized 20 h later, and bacterial burdens (CFU per organ) were determined. MacConkey agar plates with specific antibiotics were used for the cultivation of ETEC 6. To estimate survival, ETEC 6 ( $10^9$  CFU) was administered orally by gavage to pups, and the condition of the mice was closely monitored. A moribund condition was recorded as the experimental end point, and survival (defined as the percentage of animals that were alive compared to those that were moribund or dead 20 h after challenge) was recorded. In the systemic infection model, ETEC 6 ( $10^7$  CFU) was administered intraperitoneally to 10–12-day-old pups, and the condition of the mice was closely monitored. A moribund condition was recorded as the experimental end point, and the survival at 3 days after infection was recorded. Moribund 6–7-day-old animals were defined as those that were grey rather than pink in colour and

were not responsive to manual stimulation; older mice (older than 10 days of age) were defined as moribund if they displayed abnormal posture, rough hair coat, exudate around eyes and/or nose, skin lesions, abnormal breathing, difficulty with ambulation, low food and water intake or self-mutilation.

### Isolation of mouse-gut commensal Enterobacteriaceae bacteria

Homogenates of small intestine were plated on MacConkey agar plates without antibiotics and incubated aerobically at 37 °C overnight. Colonies were purified and DNA was extracted. The 16S rRNA gene was amplified by PCR and sequenced with 27F and 1492R primers<sup>43</sup>.

### RNA sequencing

Illumina sequencing libraries were built using the Ovation RNA-Seq System V2 (NuGEN) according to the manufacturer's instructions, and were submitted to the Harvard Biopolymers Facility for sequencing on the Illumina NextSeq 500, resulting in 287 million high-quality 50-nucleotide paired-end reads. Differential expression analysis was performed with the Bioconductor package DESeq2<sup>44</sup>.

### Total IgG and IgA enzyme-linked immunosorbent assays

Sera and mucosal antibodies of the neonates were measured with a mouse IgG enzyme-linked immunosorbent assay (ELISA) kit (Abcam, ab157719) and a mouse IgA ELISA kit (Abcam, ab157717). Serum samples were diluted in the 1:20,000–1:40,000 range for IgG detection. Mucosal samples—from either the small intestine or the colon—were homogenized in 1 ml of PBS and centrifuged. Only supernatants were used for ELISA. For measurement of faecal antibodies, faecal pellets were weighed and resuspended as 100 mg ml<sup>-1</sup> stock solutions in PBS buffer before further dilution for ELISA. Results were read with a BioTek Synergy HT Multi Mode Microplate Reader at OD<sub>450</sub>. For absorption assays, formalin-fixed commensal bacteria or *Pantoea* cells ( $10^8$  CFU) were incubated with 100  $\mu$ l sera, and after 1 h bacterial cells were removed by centrifugation. Absorbed serum samples were diluted and used for ELISA as described below.

### ETEC cross-reactivity ELISA

The cross-reactivity of mouse serum antibodies with ETEC 6 cells was assessed by whole-cell ELISA. In brief, ETEC 6 bacterial cells were treated with 0.5% formalin at room temperature for 2 h and then washed twice with sodium-carbonate-coating buffer. About  $10^8$  CFU per 100  $\mu$ l of fixed ETEC 6 cells in coating buffer were added to each well of a NUNC Maxisorp ELISA plate (Thermo Fisher, 44-2404-21) and then incubated overnight at 4 °C. Wells were washed with PBST (PBS and 0.05% Tween-20) and blocked with 5% nonfat milk in PBST buffer for 2 h at room temperature. Next, 2% nonfat milk in PBST was used for serum dilutions; the addition of 50  $\mu$ l of diluted serum to each well (as replicates) was followed by incubation at room temperature for 1 h. The following secondary antibodies were used at a dilution of 1:2,000: anti-mouse immunoglobulin-HRP (SouthernBiotech, 1010-05) or anti-mouse IgG-HRP (SouthernBiotech, 1030-05). Super Aqua Blue substrate (Thermo Fisher, 00-4203-58) was used for colour detection. Titres of antibody to ETEC 6 were read with a BioTek Synergy HT Multi Mode Microplate Reader at OD<sub>405</sub>.

### Construction of the ETEC-GFP strain

The plasmid pUC18T-mini-Tn7T-Tp-gfpmut3 was electronically transformed into ETEC 6 competent cells. The successful transformant was selected, confirmed to be positive for GFP by PCR as well as by flow cytometry, and designated ETEC-GFP.

### Detection of antibody deposition on in vivo-recovered ETEC cells

To analyse the IgG and IgA coating of ETEC 6 bacteria ex vivo, mice infected with ETEC-GFP were euthanized 18–20 h after infection. Small-intestine contents were scraped off, washed and filtered (filter

pore size, 5 µm; Pall Acrodisc, 4650) for recovery of bacteria. Faecal bacteria were resuspended in PBS with a cocktail of protease inhibitors (Roche, 11873580001) and incubated with shaking at 37 °C for 5–10 min to facilitate GFP maturation and detection by flow cytometry. Faecal bacteria were blocked with 2% BSA in PBS buffer and stained with diluted (1:100) anti-mouse IgG-647 (Biolegend, 405322) or anti-mouse IgA-APC (SouthernBiotech, 1165-11). Isotype controls were Alexa Fluor 647 goat IgG (Biolegend, 403006) and rat IgG1-APC (SouthernBiotech, 0116-11). Stained bacteria were washed with PBS and analysed by MACsquant (Miltenyi Biotec). Data were analysed with FlowJo software (Tree Star).

## Exogenous antibody supplementation in $\mu MT^{-/-}$ pups

The breeding of two  $\mu MT^{-/-}$  female mice was synchronized to generate two litters of pups born within a 12-h time frame. Purified SPF mouse IgG (12 mg; mu-003-C.05, ImmunoReagents) was injected intraperitoneally into one pregnant  $\mu MT^{-/-}$  female at gestation day 18 and again at postpartum day 2. The resulting two litters of  $\mu MT^{-/-}$  pups were used for ETEC 6 infection at 1 week of age.

## Cross-fostering experiment

The breeding of a  $\mu MT^{-/-}$  female with a  $\mu MT^{-/-}$  male and the breeding of a  $\mu MT^{-/-}$  female with a  $\mu MT^{-/-}$  male were synchronized to generate pups born on the same day, for subsequent cross-fostering. After 1 week of cross-fostering, pups were used for ETEC 6 infection experiments. Serum and mucosal samples of infected pups were collected for measurement of IgG and IgA titres.

## Commensal immunization

Commensal species of Enterobacteriaceae were isolated from SPF mice. One was identified as a *Pantoea* species (referred to as *Pantoea* 1). This strain was grown in LB broth to an OD<sub>600</sub> of 1.0; cells were then collected by centrifugation and treated with 1% formalin for 1 h before three washes with PBS buffer. Formalin-fixed *Pantoea* 1 (10<sup>7</sup> CFU in 100 µl of PBS) was injected intraperitoneally into mice for priming. After 3 weeks, 10<sup>7</sup> CFU of formalin-fixed *Pantoea* 1 was again injected intraperitoneally as an immunological boost. Sera were collected from 2-week-old pups for antibody titre determination and used for immunoblotting analysis.

## Immunoblotting of bacterial lysates with immunized serum

After growth of ETEC 6, *Salmonella* and *Pantoea* 1 isolates in LB broth, bacteria (10<sup>8</sup> CFU) were collected by centrifugation, lysed with lysis buffer, and run on NuPAGE 4–12% Bis-Tris protein gels (Invitrogen, NP0335BOX) at 180 V. Separated products were transferred to nitrocellulose membranes iBlot2 NC mini stacks (Invitrogen, IB23002) with an iBlot transfer device (Invitrogen, IB21001). The nitrocellulose membranes were reacted with immunized mouse serum at a dilution of 1:500 and then blotted with 1:10,000 diluted IRDye680RD goat anti-mouse

IgG secondary antibody (LI-COR, 926-68070). Images were taken with an Odyssey Imaging system (LI-COR Biosciences).

## Pronase treatment of bacterial lysates

ETEC 6, *Enterobacter* and *Pantoea* 1 were grown in LB broth to an OD<sub>600</sub> of 1.0, collected, washed three times with PBS buffer and resuspended in half volume of original bacterial culture. Bacterial suspensions were lysed three times (15-s duration) with a Branson Ultrasonics Probe Sonicator. Pronase was added to bacterial lysates to final concentrations of 0, 0.2, 1 and 2 mg ml<sup>-1</sup>, with subsequent incubation at 42 °C for 1 h. The digested bacterial lysates were used for immunoblotting analysis.

## Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

## Data availability

16S rRNA gene profiling data and the ETEC 6 genome are available in the NCBI database under BioProject PRJNA577743 and BioSample SAMN12263012, respectively.

- Caporaso, J. G. et al. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc. Natl Acad. Sci. USA* **108**, 4516–4522 (2011).
- Bolyen, E. et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **37**, 852–857 (2019).
- Suzuki, M. T. & Giovannoni, S. J. Bias caused by template annealing in the amplification of mixtures of 16S rRNA genes by PCR. *Appl. Environ. Microbiol.* **62**, 625–630 (1996).
- Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

**Acknowledgements** We thank all of the staff in our animal facility for their support in animal husbandry; all Kasper and Mekalanos laboratory members for comments; Y. Chen for technical help; and F. Qadri for providing the ETEC 6 strain. The study was supported by NIH NIAID Center of Excellence for Translational Research Grant 1U19 AI109764 awarded to D.L.K. and AI01845 awarded to J.J.M. G.S. was supported in part by funding from the European Union's Horizon 2020 Research and Innovation Programme under Marie Skłodowska Curie Grant Agreement 661138.

**Author contributions** W. Zheng, W. Zhao, J.J.M. and D.L.K. conceived and designed the study and wrote and revised the manuscript. W. Zheng performed all mouse breeding, genotyping and fostering experiments, ELISA and biochemistry experiments. W. Zhao and W. Zheng conducted mouse challenge experiments and immuno-analysis of antibody specificities. M.W. performed the 16S rRNA gene-sequencing study of the mouse gut microbiota. F.C. conducted the mouse transcriptome analysis. X. Song helped to maintain the mouse line with genotyping and provided insights on mouse breeding, western blot and other techniques. X. Sun, F.G., G.S. and S.O. provided discussions during the generation of the research.

**Competing interests** The authors declare no competing interests.

## Additional information

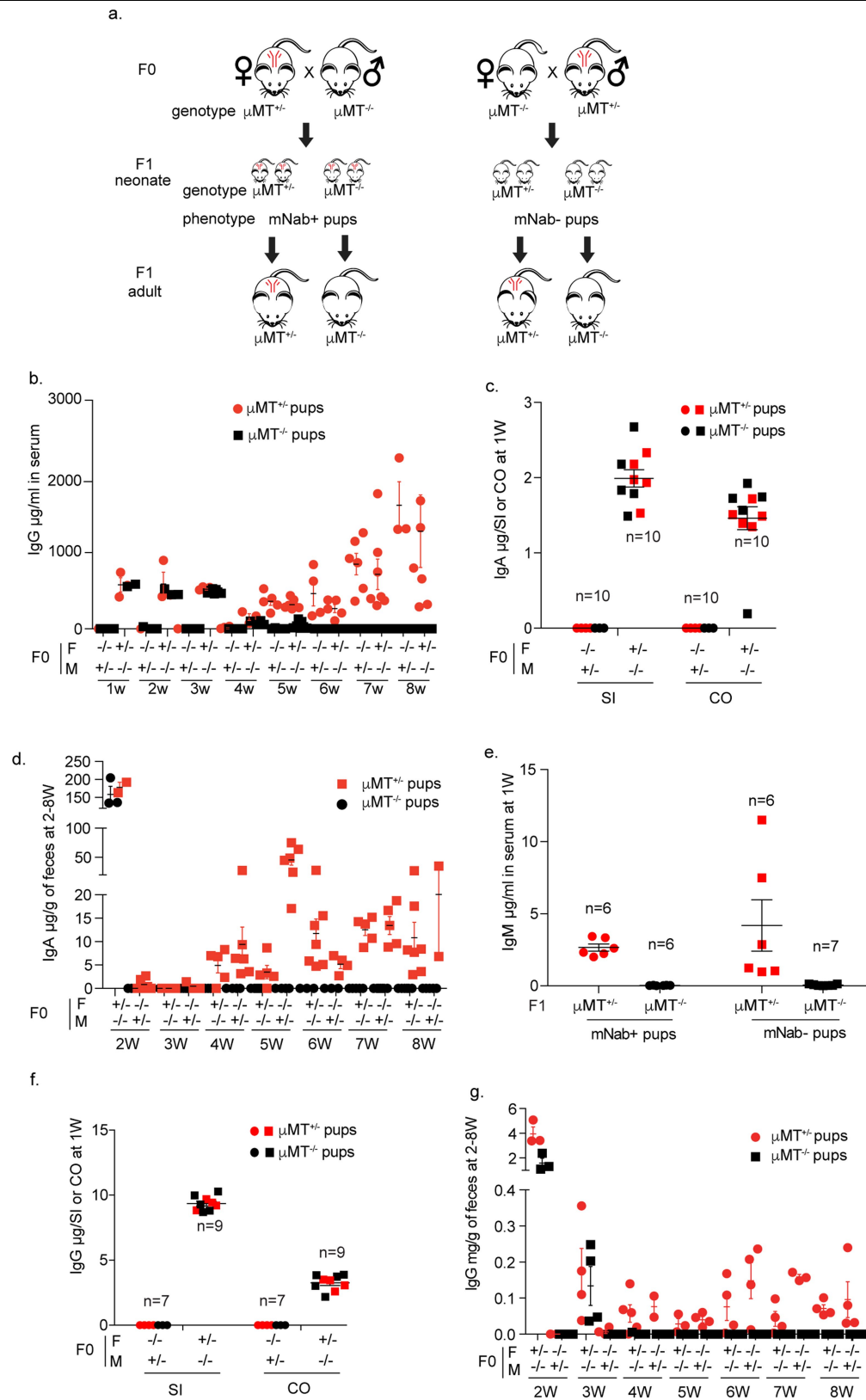
**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41586-019-1898-4>.

**Correspondence and requests for materials** should be addressed to D.L.K.

**Peer review information** Nature thanks Duane Wesemann and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>.





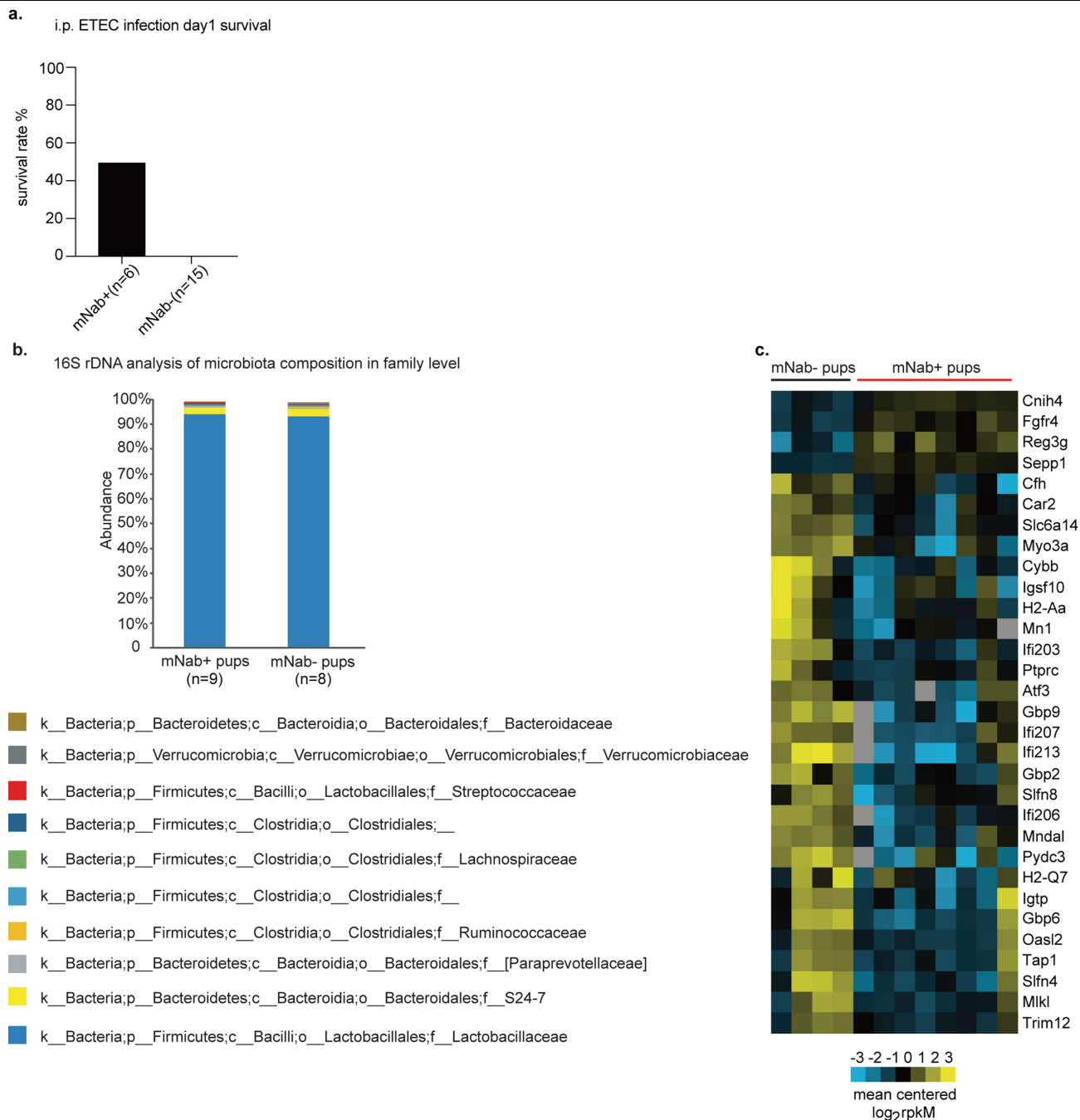
**Extended Data Fig. 1** | See next page for caption.

## Extended Data Fig. 1 | Persistence and development of maternal antibodies.

The genotypes of mating pairs are indicated under the large mouse cartoons; the small mouse cartoons represent neonates born to the indicated dams and their genotypes. Red symbols denote the presence in neonates of antibodies that either were acquired transplacentally from  $\mu MT^{-/-}$  mothers or, in the case of  $\mu MT^{-/-}$  pups, were generated endogenously after 4 weeks of age.

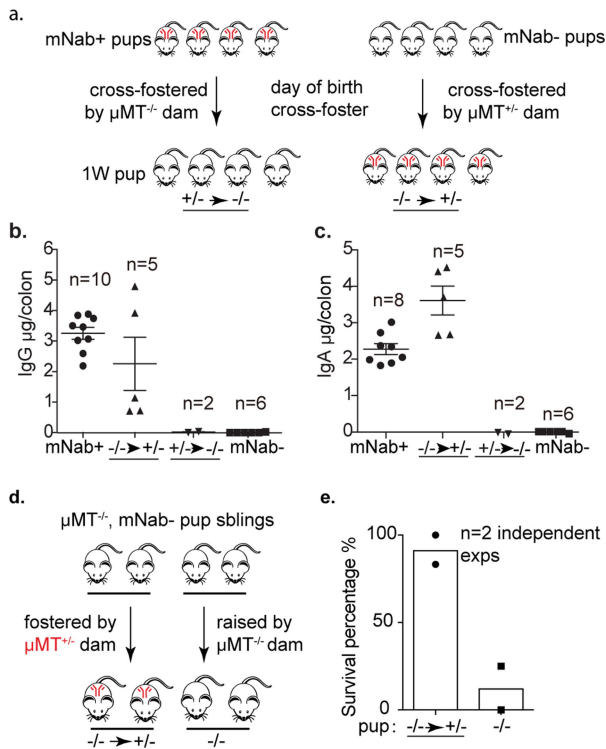
**a**, Reciprocal breeding scheme to study maternal antibody persistence and development. **b**, Serum IgG concentration in 1–8-week-old pups. Data are shown as  $\mu\text{g ml}^{-1}$ .  $n = 5$ –15 mice in each breeding group for every week of 1–8 weeks. Further details are provided in the Source Data. **c**, IgA concentrations in small-intestine (SI) and colon (CO) homogenates from

1-week-old pups. Data are shown as  $\mu\text{g}$  per small intestine or colon. **d**, Faecal IgA concentration in 2–8-week-old pups. Data are shown as  $\mu\text{g}$  per g of faeces.  $n = 6$ –13 mice in each breeding group for every week of 2–8 weeks. Further details are provided in the Source Data. **e**, Serum IgM concentration in 1-week-old pups. Data are shown as  $\mu\text{g ml}^{-1}$ . **f**, IgG concentration in small-intestine and colon homogenates from 1-week-old pups. Data are shown as  $\mu\text{g}$  per small intestine or colon. **g**, Faecal IgG concentration in 2–8-week-old pups. Data are shown as  $\mu\text{g}$  per g of faeces.  $n = 5$ –9 mice in each breeding group for every week of 2–8 weeks. Further details are provided in the Source Data. **b–g**, Data are mean  $\pm$  s.e.m. Specific  $n$  numbers are shown in the figure.



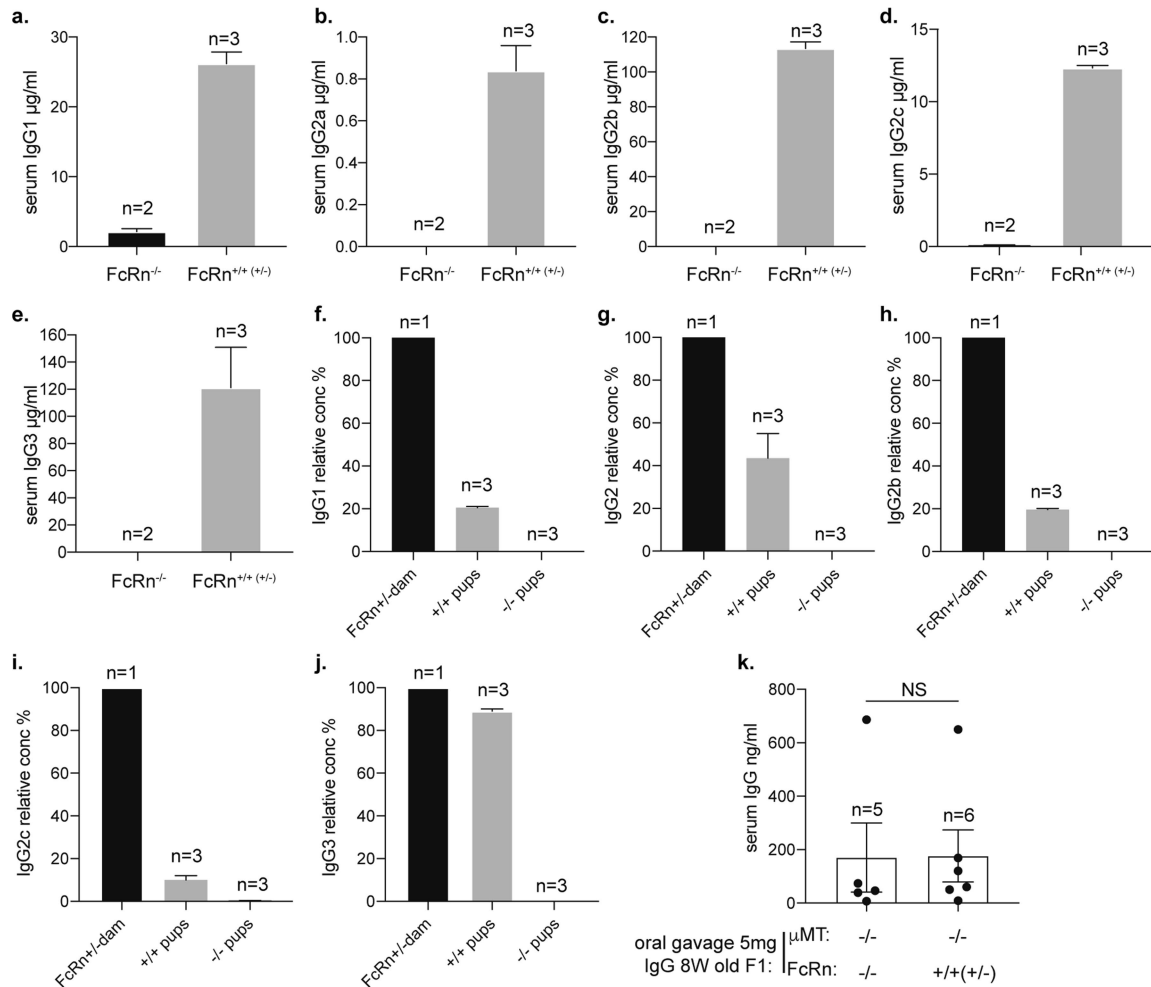
**Extended Data Fig. 2 | Comparison of mNab<sup>+</sup> and mNab<sup>-</sup> pups.** **a**, Survival of 2-week-old mNab<sup>+</sup> and mNab<sup>-</sup> pups on day 1 after intraperitoneal challenge with 10<sup>7</sup> CFU of ETEC 6. mNab<sup>+</sup> group, *n* = 6 pups; mNab<sup>-</sup> group, *n* = 15 pups. **b**, 16S rDNA gene analysis of the composition of the microbiota in 1-week-old reciprocally bred pups. Data are the average of 8 or 9 individual pups from each

group; mNab<sup>+</sup> group, *n* = 9 pups; mNab<sup>-</sup> group, *n* = 8 pups. **c**, Transcriptome analysis of small intestines of ETEC-6-infected mNab<sup>+</sup> and mNab<sup>-</sup> pups using RNA sequencing. *n* = 8 mNab<sup>+</sup> pups; *n* = 4 mNab<sup>-</sup> pups. Specific *n* numbers are shown in the figure.



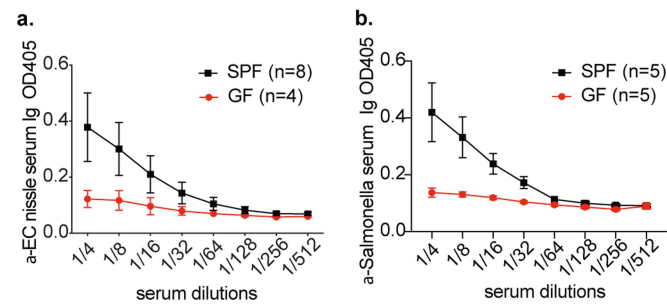
**Extended Data Fig. 3 | Comparison of pups from the cross-fostering experiment.** For all panels, cross-fostering of neonates is denoted by horizontal arrows that provide the genotype of the pup followed by the genotype of the fostering dam. **a**, Cross-fostering experimental scheme. The genotypes of dams are indicated under the large mouse cartoons; the small mouse cartoons represent neonates that are born to those dams and have the same genotype as their mothers. Thicker arrows define the mother that fostered the indicated neonates. Red symbols denote the presence in neonates of antibodies that were acquired transplacentally from their  $\mu\text{MT}^{-/-}$  mothers or, in the case of  $\mu\text{MT}^{-/-}$  pups, from a  $\mu\text{MT}^{-/-}$  fostering dam. **b**, Colon IgG concentration of 1-week-old cross-fostered pups. Data are shown as  $\mu\text{g}$  per colon. **c**, Colon IgA concentration of 1-week-old cross-fostered pups. Data are shown as  $\mu\text{g}$  per colon. **d**, Fostering scheme of  $\mu\text{MT}^{-/-}$  pups cross-fostered by a  $\mu\text{MT}^{-/-}$  dam. **e**, Survival of fostered  $\mu\text{MT}^{-/-}$  pups at 20-h after ETEC infection. In the first experiment,  $n = 5$   $\mu\text{MT}^{-/-}$  pups were fostered by  $\mu\text{MT}^{-/-}$  dams;  $n = 5$   $\mu\text{MT}^{-/-}$  pups were fostered by  $\mu\text{MT}^{-/-}$  dams. In the second experiment,  $n = 6$   $\mu\text{MT}^{-/-}$  pups were cross-fostered by  $\mu\text{MT}^{-/-}$  dams;  $n = 8$   $\mu\text{MT}^{-/-}$  pups were fostered by  $\mu\text{MT}^{-/-}$  dams. **b**, **c**, Data are mean  $\pm$  s.e.m. Specific  $n$  numbers are shown in the figure.



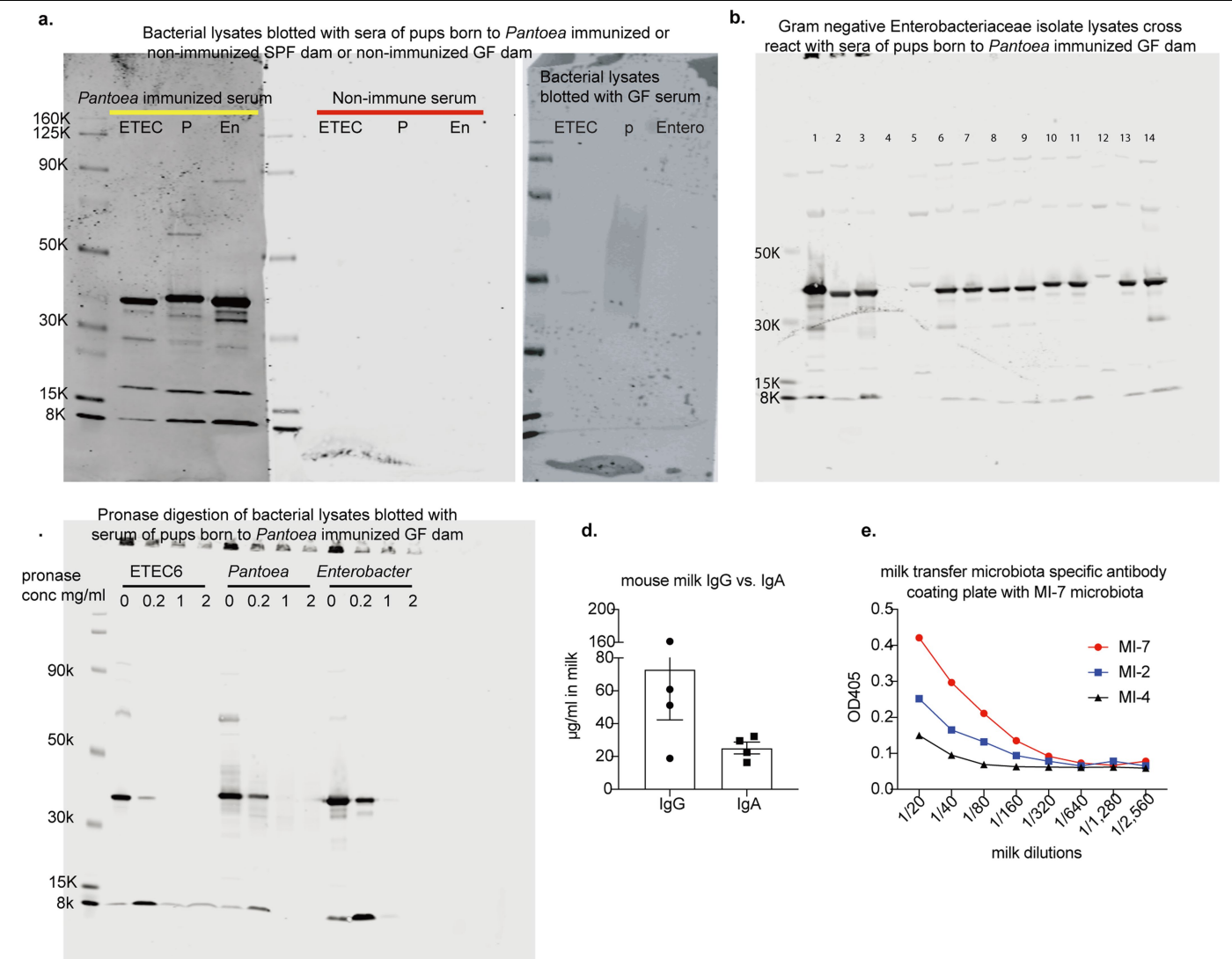


**Extended Data Fig. 4 | Relative concentrations of all subclasses of IgG between dams and pups.** **a–e.** Serum IgG subclass concentrations in  $\mu\text{MT}^{-/-}\text{FcRn}^{+/+}$  (or  $\text{FcRn}^{+/+}$ ) and  $\mu\text{MT}^{-/-}\text{FcRn}^{-/-}$  pups fostered by  $\mu\text{MT}^{+/+}$  dams for 1 week. **f–j.** Relative concentrations of all subclasses of IgG between dam and

pups. **k.** Adult (8-week-old) mice were orally gavaged with 5 mg of IgG, and serum IgG concentrations were quantified as  $\text{ng ml}^{-1}$ . NS, no significant difference; calculated using a Mann–Whitney *U*-test. Specific *n* numbers are shown in the figure.

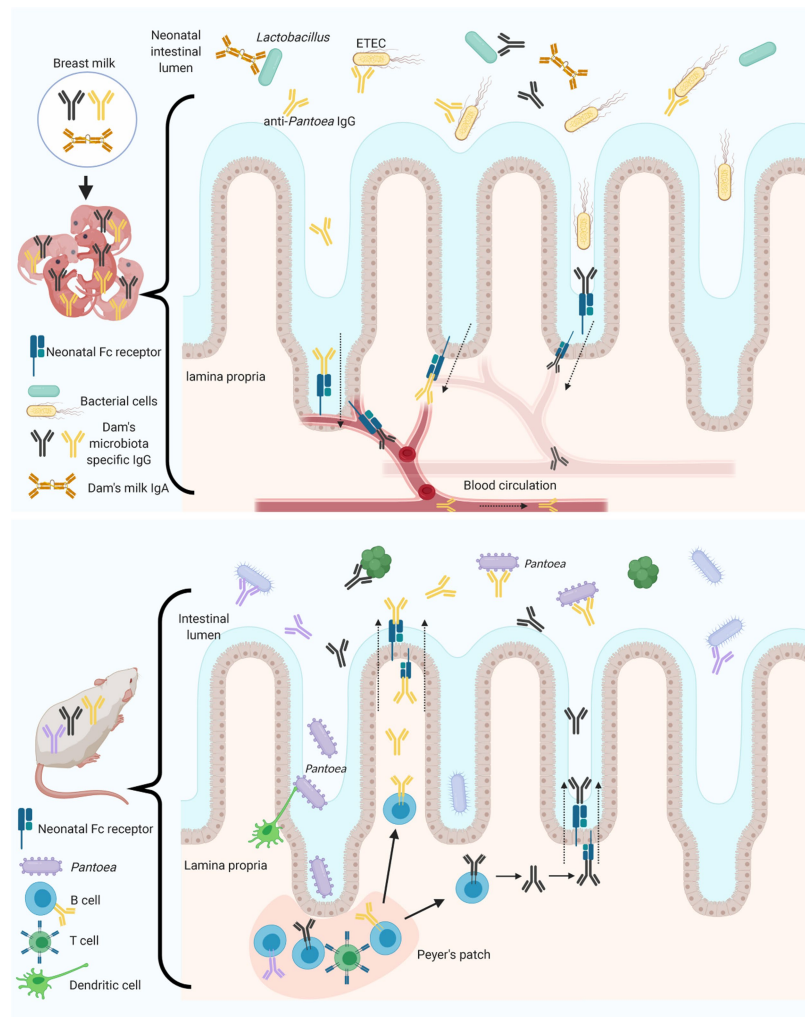


**Extended Data Fig. 5 | Serum from conventionally colonized (SPF) mice broadly recognizes human commensal bacteria and other enteric pathogens. a,** Total immunoglobulin titres against *E. coli* strain Nissle 1917 in germ-free and SPF mouse serum. **b,** Total immunoglobulin titres against *Salmonella typhimurium* in germ-free and SPF mouse serum. Specific *n* numbers are shown in the figure.



**Extended Data Fig. 6 | Characteristics of commensal-immunized and unimmunized serum.** **a.** Western blot analysis of serum IgG from pups born to *Pantoea*-1-immunized SPF mice shows epitopes of ETEC 6, *Pantoea* 1 and *Enterobacter*, similar in size to those in serum from pups born to *Pantoea*-1-immunized germ-free mice. **b.** Serum IgG of pups born to *Pantoea*-immunized germ-free dams cross-reacts with different Enterobacteriaceae isolates from different facilities. 1, Harvard SGM *Pantoea*; 2, ETEC 6; 3, Harvard SGM *Enterobacter*; 4, *Bacteroides fragilis* NCTC9343; 5, Charles River B6 *Proteus mirabilis*; 6, Charles River B6 *E. coli* isolate 1; 7, Charles River B6 *E. coli* isolate 2;

8, Charles River CD1 *E. coli* isolate 1; 9, Charles River CD1 *E. coli* isolate 2; 10, Taconic B6 *E. coli* isolate 1; 11, Taconic B6 *E. coli* isolate 2; 12, Taconic B6 *Proteus* isolate; 13, Taconic B6 *Enterobacter* isolate; 14, *C. rodentium*. **c.** Pronase-treated bacterial lysates blotted with serum IgG of pups born to *Pantoea*-immunized germ-free dams. The concentrations of pronase are specified in the figure. **a–c.** Proteins were detected using a goat anti-mouse IgG antibody. For gel source data, see Supplementary Fig. 1. **d.** Mouse milk IgG and IgA concentrations. Data are shown in  $\mu\text{g ml}^{-1}$ . **e.** Mouse milk IgG titre against microbiota. Each line represents an independent mouse. **d.** Data are mean  $\pm$  s.e.m.



**Extended Data Fig. 7 | Schematic summary of the findings in this study.** Top, mNabs induced by commensal microbiota in dams were transferred to neonates through the breast milk. Cross-reacting mNabs (especially IgG antibodies) were detected that bound to the pathogenic, non-indigenous bacterial species ETEC and correlated with protection against disease in pups challenged with ETEC. IgG antibodies were also shown to be transported from the milk to the bloodstream of pups by a process that we call IgG retro-transport. Bottom, mNabs react with many commensal species and among

them an Enterobacteriaceae isolate (*Pantoea*) was found to induce antibodies that cross-react with ETEC. The immunogenicity of this commensal species is hypothesized to be a result of local antigen-sampling processes that involve dendritic cells and uptake by Peyer's patch germinal centres. This ultimately leads to the induction of high-affinity IgGs directed against a *Pantoea* antigen that cross-reacts with ETEC. IgG was also shown to be transported from the blood stream to the intestinal lumen by FcRn in adult mice. Illustrations were created with BioRender (<https://biorender.com/>).



## Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- ☐ ☒ The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- ☐ ☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- ☐ ☒ The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- ☒ ☐ A description of all covariates tested
- ☐ ☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- ☐ ☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- ☐ ☒ For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give  $P$  values as exact values whenever suitable.*
- ☒ ☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- ☒ ☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- ☒ ☐ Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

Data collection

No data collection software was used

Data analysis

1. GraphPad Prism 8 was used to analyze bacterial burden, mouse survival, antibody titer absorption assay, ELISA assays results.
2. 16S rDNA sequence reads were analyzed with QIIME and QIIME2 pipelines.
3. RNAseq data differential expression analysis was performed using Bio-conductor package DESeq2.
4. Flow cytometry data were analyzed by flow Jo software (Tree Star v10.6.0).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

### Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All data generated or analysed during this study are included in this published article (and its supplementary information files). Any other raw data pertaining to this study are available from the corresponding authors upon reasonable request.

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Statistical methods were not used to determine sample size. It's impossible to predict the magnitude of experimental variation between animals based on our current knowledge in this exploratory study. The group sizes (at least three animals per treatment group) represents the minimum number animals needed to reach statistical significance ( $p < 0.05$ ) between experimental groups.
Data exclusions	No data was excluded.
Replication	Most experiments were repeated between 2-6 times and each experiment involved at least 3 mice. Experiments results were robust and reproducible.
Randomization	Due to the nature of the experiment design, randomization of animal was not relevant to our study.
Blinding	The investigators were blinded to group allocation during data collection and analysis, because one investigator collected mice bacterial burden or survival, and the other person genotyping of all the mice. Two persons put the results together after they finish each analysis.

## Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	Briefly describe the study type including whether data are quantitative, qualitative, or mixed-methods (e.g. qualitative cross-sectional, quantitative experimental, mixed-methods case study).
Research sample	State the research sample (e.g. Harvard university undergraduates, villagers in rural India) and provide relevant demographic information (e.g. age, sex) and indicate whether the sample is representative. Provide a rationale for the study sample chosen. For studies involving existing datasets, please describe the dataset and source.
Sampling strategy	Describe the sampling procedure (e.g. random, snowball, stratified, convenience). Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient. For qualitative data, please indicate whether data saturation was considered, and what criteria were used to decide that no further sampling was needed.
Data collection	Provide details about the data collection procedure, including the instruments or devices used to record the data (e.g. pen and paper, computer, eye tracker, video or audio equipment) whether anyone was present besides the participant(s) and the researcher, and whether the researcher was blind to experimental condition and/or the study hypothesis during data collection.
Timing	Indicate the start and stop dates of data collection. If there is a gap between collection periods, state the dates for each sample cohort.
Data exclusions	If no data were excluded from the analyses, state so OR if data were excluded, provide the exact number of exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.
Non-participation	State how many participants dropped out/declined participation and the reason(s) given OR provide response rate OR state that no participants dropped out/declined participation.
Randomization	If participants were not allocated into experimental groups, state so OR describe how participants were allocated to groups, and if allocation was not random, describe how covariates were controlled.

## Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	Briefly describe the study. For quantitative data include treatment factors and interactions, design structure (e.g. factorial, nested, hierarchical), nature and number of experimental units and replicates.
Research sample	Describe the research sample (e.g. a group of tagged <i>Passer domesticus</i> , all <i>Stenocereus thurberi</i> within Organ Pipe Cactus National Monument), and provide a rationale for the sample choice. When relevant, describe the organism taxa, source, sex, age range and

any manipulations. State what population the sample is meant to represent when applicable. For studies involving existing datasets, describe the data and its source.

#### Sampling strategy

Note the sampling procedure. Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient.

#### Data collection

Describe the data collection procedure, including who recorded the data and how.

#### Timing and spatial scale

Indicate the start and stop dates of data collection, noting the frequency and periodicity of sampling and providing a rationale for these choices. If there is a gap between collection periods, state the dates for each sample cohort. Specify the spatial scale from which the data are taken

#### Data exclusions

If no data were excluded from the analyses, state so OR if data were excluded, describe the exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.

#### Reproducibility

Describe the measures taken to verify the reproducibility of experimental findings. For each experiment, note whether any attempts to repeat the experiment failed OR state that all attempts to repeat the experiment were successful.

#### Randomization

Describe how samples/organisms/participants were allocated into groups. If allocation was not random, describe how covariates were controlled. If this is not relevant to your study, explain why.

#### Blinding

Describe the extent of blinding used during data acquisition and analysis. If blinding was not possible, describe why OR explain why blinding was not relevant to your study.

Did the study involve field work? ☐ Yes ☐ No

## Field work, collection and transport

#### Field conditions

Describe the study conditions for field work, providing relevant parameters (e.g. temperature, rainfall).

#### Location

State the location of the sampling or experiment, providing relevant parameters (e.g. latitude and longitude, elevation, water depth).

#### Access and import/export

Describe the efforts you have made to access habitats and to collect and import/export your samples in a responsible manner and in compliance with local, national and international laws, noting any permits that were obtained (give the name of the issuing authority, the date of issue, and any identifying information).

#### Disturbance

Describe any disturbance caused by the study and how it was minimized.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

- |                                     |   |
|-------------------------------------|---|
| n/a                                 | Involved in the study   |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> Antibodies                  |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Eukaryotic cell lines                  |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Palaeontology                          |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> Animals and other organisms |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Human research participants            |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Clinical data                          |

### Methods

- |                                     |  |
|-------------------------------------|--|
| n/a                                 | Involved in the study                              |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> ChIP-seq                  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> Flow cytometry |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> MRI-based neuroimaging    |

## Antibodies

#### Antibodies used

The following antibodies were used: Name / Clone / Cat# / Dilution / Manufacture  
 Goat-anti-mouse-Ig-HRP / Polyclonal / Cat.1010-05 / 1:2000 / Southern Biotech  
 Goat anti-mouse-IgG-HRP / Polyclonal / Cat.1030-05 / 1:2000 / Southern Biotech  
 AlexaFluor@647Goat anti-mouse IgG / Poly4053 / Cat. 405322 / 1:100 / Biolegend  
 AlexaFluor@647 Goat IgG / Poly24030 / Cat. 403006 / 1:100 / Biolegend  
 Rat anti-mouse IgA-APC / 11-44-2 / Cat. 1165-11 / 1:100 / Southern Biotech  
 Rat IgG1-APC / KLH/G1-2-2 / Cat. 0116-11 / 1:100 / Southern Biotech  
 IRDye680RD goat anti-mouse IgG secondary antibody / Cat. 926-68070 / 1:10000 / LI-COR

The following ELISA kits and antibody were used:  
 Mouse IgG ELISA kit (Abcam Cat. ab157719)  
 Mouse IgA ELISA kit (Abcam Cat. ab157717)

Mouse IgM ELISA Kit (Abcam Cat. ab133047)  
 Mouse IgG1 ELISA Kit (Abcam Cat. ab133045)  
 Mouse IgG2a ELISA Kit (Abcam, Cat. ab133046)  
 Mouse IgG2b ELISA Kit (Abcam, Cat. ab136941)  
 Mouse IgG2c ELISA Kit (Abcam, Cat. ab157720)  
 Mouse IgG3 ELISA Kit (Abcam, Cat. ab157721)  
 Purified SPF mouse IgG / Cat. mu-003-C.05 / ImmunoReagents

## Validation

All antibodies and ELISA kits were commercially available in Abcam, Southern Biotech, Biolegend, ImmunoReagents or LI-COR. The validation data is available on their websites.

## Eukaryotic cell lines

Policy information about [cell lines](#)

## Cell line source(s)

*State the source of each cell line used.*

## Authentication

*Describe the authentication procedures for each cell line used OR declare that none of the cell lines used were authenticated.*

## Mycoplasma contamination

*Confirm that all cell lines tested negative for mycoplasma contamination OR describe the results of the testing for mycoplasma contamination OR declare that the cell lines were not tested for mycoplasma contamination.*

Commonly misidentified lines  
(See [ICLAC](#) register)

*Name any commonly misidentified cell lines used in the study and provide a rationale for their use.*

## Palaeontology

## Specimen provenance

*Provide provenance information for specimens and describe permits that were obtained for the work (including the name of the issuing authority, the date of issue, and any identifying information).*

## Specimen deposition

*Indicate where the specimens have been deposited to permit free access by other researchers.*

## Dating methods

*If new dates are provided, describe how they were obtained (e.g. collection, storage, sample pretreatment and measurement), where they were obtained (i.e. lab name), the calibration program and the protocol for quality assurance OR state that no new dates are provided.*

☐ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

## Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

## Laboratory animals

Mus Musculus, C57BL/6, neonates and pups up to 8 week old. uMT<sup>-/-</sup> mice (Stock Number: 002288), FcRn<sup>-/-</sup> mice (Stock number: 003982) and WT C57BL/6J mice (Stock number: 000664) were purchased from Jackson Laboratory and maintained at Harvard Medical School Seely Mudd Animal facility. CD1 mice (strain code: 022) and C57BL/6 mice (strain code: 027) were purchased from Charles river and maintained at Harvard Medical School Seely Mudd Animal facility. C57BL/6 mice (nomenclature: C57BL/6NTac) were purchased from Taconic and maintained at Harvard Medical School Seely Mudd Animal facility.

## Wild animals

The study did not involve wild animals.

## Field-collected samples

The study did not involve samples collected from the field.

## Ethics oversight

All animal studies were approved by IACUC of Harvard Medical School under the animal protocol IS:00000178-3.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Human research participants

Policy information about [studies involving human research participants](#)

## Population characteristics

*Describe the covariate-relevant population characteristics of the human research participants (e.g. age, gender, genotypic information, past and current diagnosis and treatment categories). If you filled out the behavioural & social sciences study design questions and have nothing to add here, write "See above."*

## Recruitment

*Describe how participants were recruited. Outline any potential self-selection bias or other biases that may be present and how these are likely to impact results.*

## Ethics oversight

*Identify the organization(s) that approved the study protocol.*

Note that full information on the approval of the study protocol must also be provided in the manuscript.



## Clinical data

Policy information about [clinical studies](#)

All manuscripts should comply with the ICMJE [guidelines for publication of clinical research](#) and a completed [CONSORT checklist](#) must be included with all submissions.

Clinical trial registration	Provide the trial registration number from ClinicalTrials.gov or an equivalent agency.
Study protocol	Note where the full trial protocol can be accessed OR if not available, explain why.
Data collection	Describe the settings and locales of data collection, noting the time periods of recruitment and data collection.
Outcomes	Describe how you pre-defined primary and secondary outcome measures and how you assessed these measures.

## ChIP-seq

### Data deposition

- ☐ Confirm that both raw and final processed data have been deposited in a public database such as [GEO](#).
- ☐ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links <i>May remain private before publication.</i>	For "Initial submission" or "Revised version" documents, provide reviewer access links. For your "Final submission" document, provide a link to the deposited data.
Files in database submission	Provide a list of all files available in the database submission.
Genome browser session (e.g. <a href="#">UCSC</a> )	Provide a link to an anonymized genome browser session for "Initial submission" and "Revised version" documents only, to enable peer review. Write "no longer applicable" for "Final submission" documents.

### Methodology

Replicates	Describe the experimental replicates, specifying number, type and replicate agreement.
Sequencing depth	Describe the sequencing depth for each experiment, providing the total number of reads, uniquely mapped reads, length of reads and whether they were paired- or single-end.
Antibodies	Describe the antibodies used for the ChIP-seq experiments; as applicable, provide supplier name, catalog number, clone name, and lot number.
Peak calling parameters	Specify the command line program and parameters used for read mapping and peak calling, including the ChIP, control and index files used.
Data quality	Describe the methods used to ensure data quality in full detail, including how many peaks are at FDR 5% and above 5-fold enrichment.
Software	Describe the software used to collect and analyze the ChIP-seq data. For custom code that has been deposited into a community repository, provide accession details.

## Flow Cytometry

### Plots

Confirm that:

- ☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- ☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- ☐ All plots are contour plots with outliers or pseudocolor plots.
- ☐ A numerical value for number of cells or percentage (with statistics) is provided.

### Methodology

Sample preparation	Mouse small intestine contents were scraped, washed and filtered through 5um filter (Pall acrodisc Cat. 4650) to recover bacteria. Fecal bacteria were resuspended in PBS with a cocktail of protease inhibitors (Roche Cat. 11873580001) and incubated with shaking at 37oC incubator for 5-10 mins to facilitate GFP protein maturation and detection on flow cytometry. Fecal bacteria were blocked with 2% BSA in PBS buffer and stained with 1:100 diluted anti-mouse IgG-647 (Biolegend Cat. 405322) and anti-mouse IgA-647 respectively. Stained bacteria were washed with PBS and analyzed by MACsquant (Miltenyi Biotec). Data were analyzed by flow Jo software (V10.6.0)
--------------------	---

Instrument	MACsquant (Miltenyi Biotec)
Software	Data were analyzed by flow Jo software (Tree Star)
Cell population abundance	Describe the abundance of the relevant cell populations within post-sort fractions, providing details on the purity of the samples and how it was determined.
Gating strategy	Describe the gating strategy used for all relevant experiments, specifying the preliminary FSC/SSC gates of the starting cell population, indicating where boundaries between "positive" and "negative" staining cell populations are defined.

☐ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

## Magnetic resonance imaging

### Experimental design

Design type	Indicate task or resting state; event-related or block design.
Design specifications	Specify the number of blocks, trials or experimental units per session and/or subject, and specify the length of each trial or block (if trials are blocked) and interval between trials.
Behavioral performance measures	State number and/or type of variables recorded (e.g. correct button press, response time) and what statistics were used to establish that the subjects were performing the task as expected (e.g. mean, range, and/or standard deviation across subjects).

### Acquisition

Imaging type(s)	Specify: functional, structural, diffusion, perfusion.
Field strength	Specify in Tesla
Sequence & imaging parameters	Specify the pulse sequence type (gradient echo, spin echo, etc.), imaging type (EPI, spiral, etc.), field of view, matrix size, slice thickness, orientation and TE/TR/flip angle.
Area of acquisition	State whether a whole brain scan was used OR define the area of acquisition, describing how the region was determined.
Diffusion MRI	<input type="checkbox"/> Used <input type="checkbox"/> Not used

### Preprocessing

Preprocessing software	Provide detail on software version and revision number and on specific parameters (model/functions, brain extraction, segmentation, smoothing kernel size, etc.).
Normalization	If data were normalized/standardized, describe the approach(es): specify linear or non-linear and define image types used for transformation OR indicate that data were not normalized and explain rationale for lack of normalization.
Normalization template	Describe the template used for normalization/transformation, specifying subject space or group standardized space (e.g. original Talairach, MNI305, ICBM152) OR indicate that the data were not normalized.
Noise and artifact removal	Describe your procedure(s) for artifact and structured noise removal, specifying motion parameters, tissue signals and physiological signals (heart rate, respiration).
Volume censoring	Define your software and/or method and criteria for volume censoring, and state the extent of such censoring.

### Statistical modeling & inference

Model type and settings	Specify type (mass univariate, multivariate, RSA, predictive, etc.) and describe essential details of the model at the first and second levels (e.g. fixed, random or mixed effects; drift or auto-correlation).
Effect(s) tested	Define precise effect in terms of the task or stimulus conditions instead of psychological concepts and indicate whether ANOVA or factorial designs were used.
Specify type of analysis:	<input type="checkbox"/> Whole brain <input type="checkbox"/> ROI-based <input type="checkbox"/> Both
Statistic type for inference (See <a href="#">Eklund et al. 2016</a> )	Specify voxel-wise or cluster-wise and report all relevant parameters for cluster-wise methods.
Correction	Describe the type of correction and how it is obtained for multiple comparisons (e.g. FWE, FDR, permutation or Monte Carlo).

## Models &amp; analysis

- n/a | Involved in the study
- ☐ ☐ Functional and/or effective connectivity
- ☐ ☐ Graph analysis
- ☐ ☐ Multivariate modeling or predictive analysis

Functional and/or effective connectivity

*Report the measures of dependence used and the model details (e.g. Pearson correlation, partial correlation, mutual information).*

Graph analysis

*Report the dependent variable and connectivity measure, specifying weighted graph or binarized graph, subject- or group-level, and the global and/or node summaries used (e.g. clustering coefficient, efficiency, etc.).*

Multivariate modeling and predictive analysis

*Specify independent variables, features extraction and dimension reduction, model, training and evaluation metrics.*

# B cells and tertiary lymphoid structures promote immunotherapy response

<https://doi.org/10.1038/s41586-019-1922-8>

Received: 5 February 2019

Accepted: 4 December 2019

Published online: 15 January 2020

Beth A. Helmink<sup>1,24\*</sup>, Sangeetha M. Reddy<sup>2,24</sup>, Jianjun Gao<sup>3,24</sup>, Shaojun Zhang<sup>4,24</sup>, Rafet Basar<sup>5,24</sup>, Rohit Thakur<sup>1</sup>, Keren Yizhak<sup>6</sup>, Moshe Sade-Feldman<sup>6,7</sup>, Jorge Blando<sup>8</sup>, Guangchun Han<sup>4</sup>, Vancheswaran Gopalakrishnan<sup>1</sup>, Yuanxin Xi<sup>9</sup>, Hao Zhao<sup>8</sup>, Rodabe N. Amaria<sup>10</sup>, Hussein A. Tawbi<sup>10</sup>, Alex P. Cogdill<sup>1</sup>, Wenbin Liu<sup>8</sup>, Valerie S. LeBleu<sup>11</sup>, Fernanda G. Kugeratski<sup>11</sup>, Sapna Patel<sup>10</sup>, Michael A. Davies<sup>10</sup>, Patrick Hwu<sup>10</sup>, Jeffrey E. Lee<sup>1</sup>, Jeffrey E. Gershenwald<sup>1</sup>, Anthony Lucci<sup>1</sup>, Reetakshi Arora<sup>4</sup>, Scott Woodman<sup>10</sup>, Emily Z. Keung<sup>1</sup>, Pierre-Olivier Gaudreau<sup>1</sup>, Alexandre Reuben<sup>12</sup>, Christine N. Spencer<sup>13</sup>, Elizabeth M. Burton<sup>1</sup>, Lauren E. Haydu<sup>1</sup>, Alexander J. Lazar<sup>4,14,15</sup>, Roberta Zapassodi<sup>16</sup>, Courtney W. Hudgens<sup>14</sup>, Deborah A. Ledesma<sup>14</sup>, SuFey Ong<sup>17</sup>, Michael Bailey<sup>17</sup>, Sarah Warren<sup>17</sup>, Disha Rao<sup>18</sup>, Oscar Krijgsman<sup>18</sup>, Elisa A. Rozeman<sup>18</sup>, Daniel Peeper<sup>18</sup>, Christian U. Blank<sup>18</sup>, Ton N. Schumacher<sup>18</sup>, Lisa H. Butterfield<sup>19</sup>, Monika A. Zelazowska<sup>20</sup>, Kevin M. McBride<sup>20</sup>, Raghu Kalluri<sup>11</sup>, James Allison<sup>8</sup>, Florent Petitprez<sup>21,22,23</sup>, Wolf Herman Fridman<sup>21,22</sup>, Catherine Sautès-Fridman<sup>21,22</sup>, Nir Hacohen<sup>6,7</sup>, Katayoun Rezvani<sup>5,25</sup>, Padmanee Sharma<sup>3,8,25</sup>, Michael T. Tetzlaff<sup>14,15,25</sup>, Linghua Wang<sup>4,25</sup> & Jennifer A. Wargo<sup>1,4,25\*</sup>

Treatment with immune checkpoint blockade (ICB) has revolutionized cancer therapy. Until now, predictive biomarkers<sup>1–10</sup> and strategies to augment clinical response have largely focused on the T cell compartment. However, other immune subsets may also contribute to anti-tumour immunity<sup>11–15</sup>, although these have been less well-studied in ICB treatment<sup>16</sup>. A previously conducted neoadjuvant ICB trial in patients with melanoma showed via targeted expression profiling<sup>17</sup> that B cell signatures were enriched in the tumours of patients who respond to treatment versus non-responding patients. To build on this, here we performed bulk RNA sequencing and found that B cell markers were the most differentially expressed genes in the tumours of responders versus non-responders. Our findings were corroborated using a computational method (MCP-counter<sup>18</sup>) to estimate the immune and stromal composition in this and two other ICB-treated cohorts (patients with melanoma and renal cell carcinoma). Histological evaluation highlighted the localization of B cells within tertiary lymphoid structures. We assessed the potential functional contributions of B cells via bulk and single-cell RNA sequencing, which demonstrate clonal expansion and unique functional states of B cells in responders. Mass cytometry showed that switched memory B cells were enriched in the tumours of responders. Together, these data provide insights into the potential role of B cells and tertiary lymphoid structures in the response to ICB treatment, with implications for the development of biomarkers and therapeutic targets.

Immunotherapy has afforded patients with melanoma and other cancers the potential for long-term survival, and we are beginning to gain insight into the mechanisms of therapeutic responses as well as biomarkers of response and resistance. Considerable progress has been made in this regard, with the identification of several validated biomarkers, particularly for ICB therapy<sup>1–10</sup>. It is clear that cytotoxic T cells have a dominant role in responses to ICB and other forms of immunotherapy; however, there is a growing appreciation of other components of the tumour microenvironment that may influence the therapeutic response—including myeloid cells and other subsets of immune cells<sup>11</sup>.

Tumour-infiltrating B cells have been identified, but their overall functional role in cancer is incompletely understood<sup>14,15,19–24</sup>—some studies suggest that they are tumour-promoting, whereas others show a positive association with improved cancer outcomes, particularly when they are found in association with organized lymphoid aggregates known as tertiary lymphoid structures (TLSs)<sup>12,13,16,25–28</sup>.

TLSs have been identified within a wide range of human cancers at all stages of disease, in primary as well as metastatic lesions, but their presence is highly variable between cancer types as well as between patients<sup>12,16</sup>. Considerable heterogeneity also exists in the cellular constituents of TLSs and their location within tumours, and



this may influence the overall effect on anti-tumour immunity and outcome<sup>12–14,16</sup>. These TLS structures are not only a surrogate marker of a brisk immune response; instead, it is thought that they actively modulate anti-tumour immune activity. In this regard, the benefit of a high CD8<sup>+</sup> T cell density within a tumour is abrogated in the absence of TLS-associated dendritic cells<sup>29</sup>. Mature TLSs exhibit evidence for the formation of germinal centres<sup>30,31</sup>, and oligoclonal B cell responses have previously been identified in cutaneous melanoma and metastases<sup>32,33</sup>, which suggests an active humoral anti-tumour response within TLSs that is driven by B cells. Notably, although preliminary evidence suggests an association between responses to ICB and the presence of B cells, the precise role of B cells—and in particular TLSs—in response to ICB remains unclear<sup>28,34</sup>.

A phase 2 clinical trial of neoadjuvant treatment with ICB in patients with high-risk resectable (clinical stage III or oligometastatic stage IV) melanoma was recently conducted to assess the safety and feasibility of this treatment in this patient population (NCT02519322)<sup>17</sup>. Notably, longitudinal tumour samples were taken in the context of therapy, and molecular and immune profiling was performed to gain insight into the mechanisms of the therapeutic response and resistance. In these studies, known and novel biomarkers of response were identified, and targeted protein expression profiling (via Nanostring Digital Spatial Profiling) revealed significantly higher expression of B cell markers in samples before treatment (baseline) and on-treatment samples of responders to ICB<sup>17</sup>.

## B cells found in the tumours of responders

To gain a deeper understanding of potential mechanisms of therapeutic response to ICB, we performed RNA sequencing (RNA-seq) in longitudinal tumour samples from this patient cohort. In these studies, significantly higher expression of B-cell-related genes such as *MZB1*, *JCHAIN* and *IGLL5* was observed in patients that respond to ICB treatment versus non-responding patients ('responders' and 'non-responders', hereafter) at baseline ( $P < 0.001$ ) with over-representation of these genes compared to T cells and other immune markers (with evaluable tumours from seven responders and nine non-responders) (Fig. 1a, b, Supplementary Tables 1, 2). Other genes that are expected to alter the function of B cells were also significantly enriched in responders versus non-responders, such as *FCRL5*, *IDO1*, *IFNG* and *BTLA*. Low tumour purity was observed in some samples, particularly in the context of an effective therapeutic response, limiting conventional analysis of RNA-seq data. To address this, we next performed a more focused investigation of the tumour immune microenvironment using the microenvironment cell populations (MCP)-counter method<sup>18</sup> on RNA-seq data in baseline and on-treatment tumour samples—focusing more specifically on immune-related genes (Supplementary Table 3), which allowed inclusion of samples with low tumour purity (10 responders and 11 non-responders at baseline, 9 responders and 11 non-responders on-treatment). In these analyses, we again observed enrichment of a B cell signature in responders versus non-responders at baseline and early on-treatment ( $P = 0.036$  and  $0.038$ , respectively). Notably, these analyses included samples from patients with nodal and extra-nodal disease with no obvious contribution based on the site of disease (Fig. 1c, Extended Data Figs. 1a, b, 2a, Supplementary Tables 4, 13), which suggests that B cell signatures were not merely related to the presence of these tumours within lymph nodes. Findings of high B cell lineage scores in responders were replicated in samples from an additional cohort of patients with melanoma treated with neoadjuvant versus adjuvant checkpoint blockade (ClinicalTrials.gov identifier NCT02437279, OpACIN-neo trial) ( $n = 12$  responders, 6 non-responders)<sup>35</sup> (Extended Data Figs. 1d, 2c, Supplementary Tables 5, 6, 13). B cell signatures alone were predictive of response in univariable analyses (odds ratio 2.6,  $P = 0.02$  for our trial, and odds ratio 2.9,  $P = 0.03$  for combined melanoma cohorts), but not in multivariable analyses when

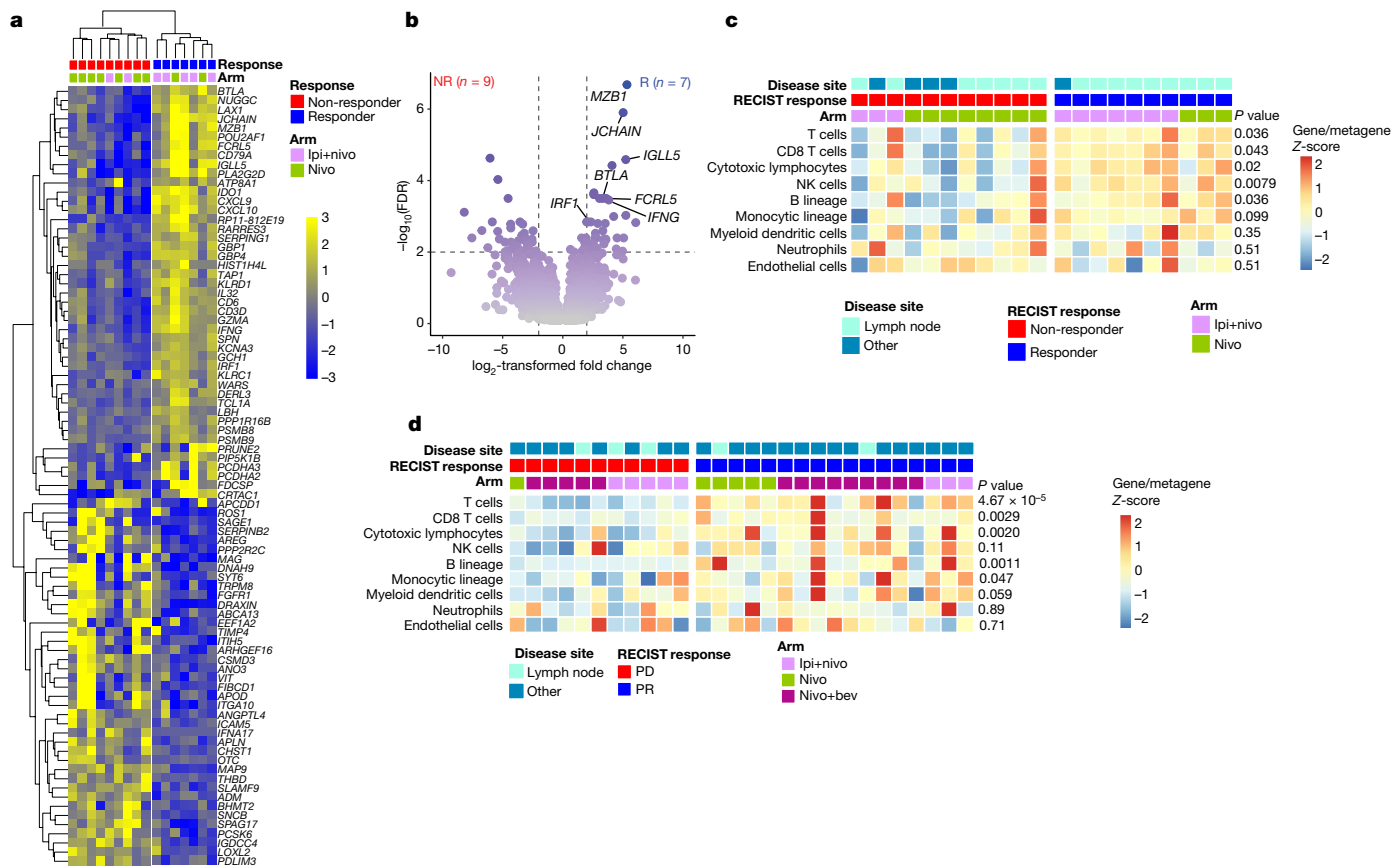
considering other components of the immune cell infiltrate, which suggests that B cells probably act together with other immune subsets and are not acting in isolation; however, these analyses were limited owing to the low sample size (Supplementary Tables 7, 8). Moreover, these findings were corroborated in translational studies of separate cohorts of patients with melanoma<sup>36</sup> and sarcoma<sup>37</sup> who were treated with ICB. B cells were not significantly associated with pathological response rates in an analogous trial of neoadjuvant-targeted therapy in patients with BRAF-mutated melanoma<sup>38</sup> (Extended Data Fig. 1e, Supplementary Table 9); however, B cells have previously been shown to be positively associated with responses to chemotherapy in other cancer types<sup>39,40</sup>.

## Similar B cell signature observed in RCC

To evaluate the validity of these findings across other cancer types, we next assessed the expression of these immune cell gene expression signatures in a pre-surgical ICB trial for patients with metastatic renal cell carcinoma (RCC) (NCT02210117, PD1 blockade monotherapy versus combined CTLA4 and PD1 blockade versus combined PD1 blockade and bevacizumab) (Supplementary Table 10). Gene expression profiling by microarray and subsequent MCP-counter analysis of baseline tumour samples was performed, demonstrating significantly higher expression of B-cell-related genes in responders versus non-responders ( $P = 0.0011$ ,  $n = 17$  responders and 11 non-responders) (Fig. 1d, Extended Data Figs. 1c, 2b, 3, Supplementary Tables 11–13). As in the case of melanoma, B cell signatures were predictive of a response in univariable analysis in the RCC cohort (odds ratio 61.2,  $P = 0.05$ ) but not multivariable analysis, again suggesting cooperative function with other immune subsets; however, sample size was again limited (Supplementary Table 14).

## B cells prognostic in TCGA analysis

On the basis of these data and existing data regarding a potential prognostic role for TLSs in melanoma and other cancer types primarily outside the context of ICB treatment<sup>18,28,41</sup>, we next assessed the expression of these immune-related genes in cutaneous melanoma from The Cancer Genome Atlas (TCGA) platform (TCGA-SKCM,  $n = 136$ )<sup>42</sup>. To this end, we applied the MCP-counter algorithm to available RNA-seq data from a subset of patients with non-recurrent stage III disease (regional lymph node or regional subcutaneous metastases), as these were most comparable to our clinical cohort. In these studies, we identified three distinct melanoma immune classes (MICs), with significantly higher expression of B cells in cluster C than in cluster A ( $P < 0.0001$ ) or cluster B ( $P < 0.0001$ ) (Extended Data Fig. 4a, Supplementary Tables 15–17). Importantly, there was no clear association of MICs with known genomic subtypes of melanoma (BRAF, NRAS, NF1 or triple wild type)<sup>42</sup> or disease site (nodal or non-nodal) (Extended Data Fig. 4a, Supplementary Table 17). Survival analyses revealed that cases in cluster C had significantly improved overall survival compared with cluster A ( $P = 0.0068$ ) (Extended Data Fig. 4b). To assess the association with B cell signatures specifically, we next compared overall survival in patients with tumours high for B cell lineage versus low, which demonstrated prolonged survival in patients with B cell-lineage-high tumours ( $P = 0.053$ ) (Extended Data Fig. 4c). Furthermore, univariable Cox proportional hazards modelling demonstrated that tumours with low infiltration of B cells had significantly increased risk of death (hazard ratio is 1.7 for B-cell-low,  $P = 0.05$ ) in comparison to the B-cell-high group (Supplementary Table 18). These data are further supported by recent analyses of the TCGA cohort that demonstrate the association of a plasmablast-like B cell signature with survival as well as increased expression of CD8A and infiltration of CD8<sup>+</sup> T cells<sup>34</sup>. Similar analyses were performed to assess the expression of immune-related genes in clear-cell RCC



**Fig. 1 | Transcriptional analysis of tumour specimens from patients with high-risk resectable melanoma and metastatic RCC treated with pre-surgical ICB. a**, Supervised hierarchical clustering of differentially expressed genes (DEG) on RNA-seq analysis by response of melanoma tumour specimens at baseline, with responder defined as having a complete or partial response by RECIST 1.1 and non-responder as having less than partial response ( $n = 9$  non-responders and 7 responders). A cut-off of gene expression fold change of  $\geq 2$  or  $\leq 0.5$  and a false discovery rate (FDR)  $q \leq 0.05$  was applied to select DEGs. Ipi, ipilimumab; nivo, nivolumab. **b**, Volcano plot depiction of DEG by response

from same cohort as in **a**. R, responders; NR, non-responders. **c**, Supervised clustering of melanoma tumour specimens by response at baseline ( $n = 11$  non-responders and 10 responders), displaying MCP-counter scores. NK cells, natural killer cells. **d**, Supervised clustering by clinical response defined as achieving a partial response (PR) according to RECIST 1.1 and non-responders as having progressive disease (PD) of RCC baseline tumour specimens ( $n = 11$  PD and 17 PR) using methodology as in **c**. P values were determined by two-sided Mann–Whitney  $U$ -test. Bev, bevacizumab.

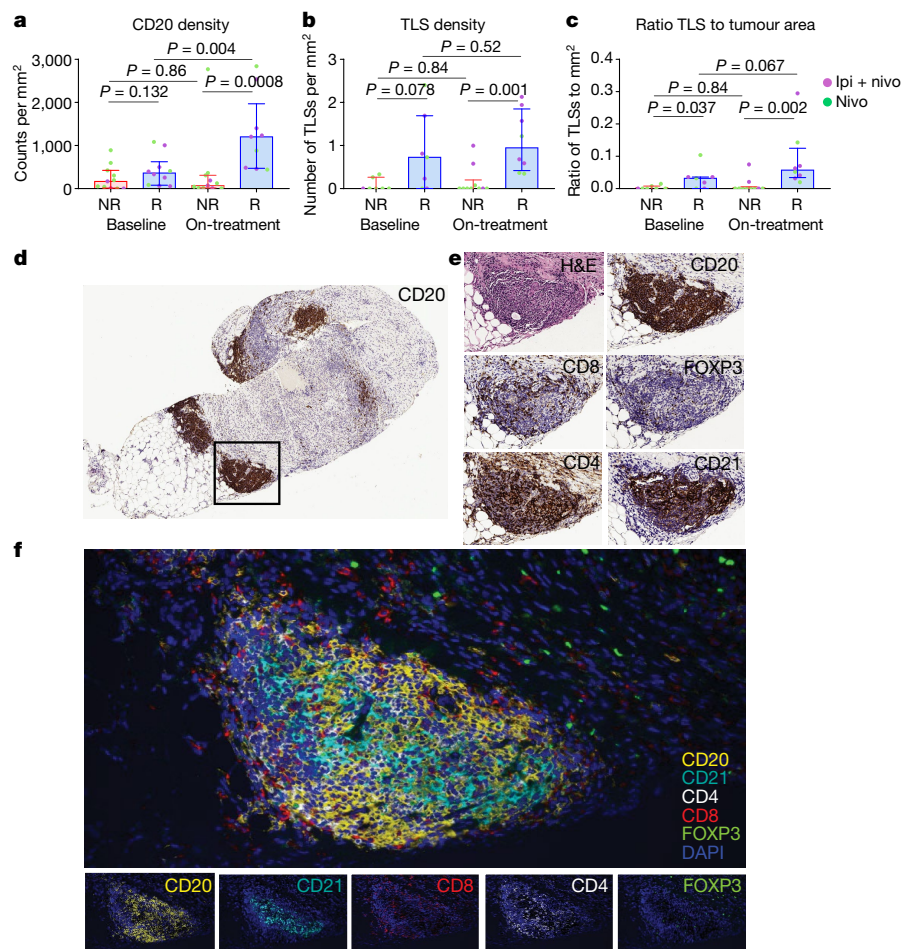
from the TCGA (TCGA-KIRC,  $n = 526$ )<sup>43</sup>. In these analyses, similar immune classes were observed; however, immune infiltration was not associated with survival in these patients ( $P = 0.24$ ) (Extended Data Fig. 4d–f, Supplementary Tables 19–21), possibly owing to the heterogeneous nature of this disease and other driving mechanisms of patient outcomes.

## B cells localized in the context of TLSs

On the basis of the results from gene expression profiling, we next assessed tumour samples histologically to gain insight into the density and distribution of B cells as well as their relationship to TLSs in patients treated with neoadjuvant ICB. The density of CD20<sup>+</sup> B cells and TLSs, and the ratio of TLSs to tumour area were higher in responders than in non-responders in our neoadjuvant melanoma cohort, particularly in early on-treatment samples ( $P = 0.0008$ ,  $P = 0.001$  and  $P = 0.002$ , respectively), although statistical significance was not reached for all the markers in the baseline samples ( $P = 0.132$ ,  $P = 0.078$  and  $P = 0.037$ , respectively) (Fig. 2a), which is consistent with previous work that suggested that assessment of early on-treatment immune infiltrate is far more predictive of the response to ICB than assessment of pre-treatment samples<sup>1</sup>. Findings between gene expression profiling and immunohistochemistry analysis were complementary, and had modest correlation as previously described<sup>18</sup> (Extended

Data Fig. 5c–e). We also found increased numbers of B-cell-related exosomes (CD20<sup>+</sup>) in the peripheral blood of responders compared with non-responders at early on-treatment time points (Extended Data Fig. 2d–j).

Notably, architectural analysis showed that CD20<sup>+</sup> B cells were localized in TLSs of tumours of responders, and were colocalized with CD4<sup>+</sup>, CD8<sup>+</sup> and FOXP3<sup>+</sup> T cells. Colocalization with CD21<sup>+</sup> follicular dendritic cells and MECA79<sup>+</sup> high endothelial venules was also shown (Fig. 2d–f, Extended Data Figs. 5a, 6a). The vast majority of evaluated TLSs in these patients represented mature secondary-follicle-like TLSs, as indicated by the presence of both CD21<sup>+</sup> follicular dendritic cells and CD23<sup>+</sup> germinal centre B cells<sup>30</sup> (Fig. 2d–f, Extended Data Figs. 5a, 6a). We identify similar mature TLSs in patients with extra-nodal metastases (Extended Data Fig. 5b), which suggests that TLSs may develop in non-nodal sites and are associated with the response to ICB treatment. Analogous immunohistochemical findings were observed in our cohort of patients with RCC treated with pre-surgical ICB, with increased infiltration of CD20<sup>+</sup> cells and TLSs density associated with response to treatment (Extended Data Fig. 6b–d); these TLSs are morphologically similar to those found in melanoma (Extended Data Fig. 6e–h). We also assessed the potential functional role of B cells and TLSs in promoting T cell responses in our cohort via additional spatial profiling analyses, and found increased markers of activation on T cells within as compared to those outside these TLSs (Extended Data Fig. 7a–c).



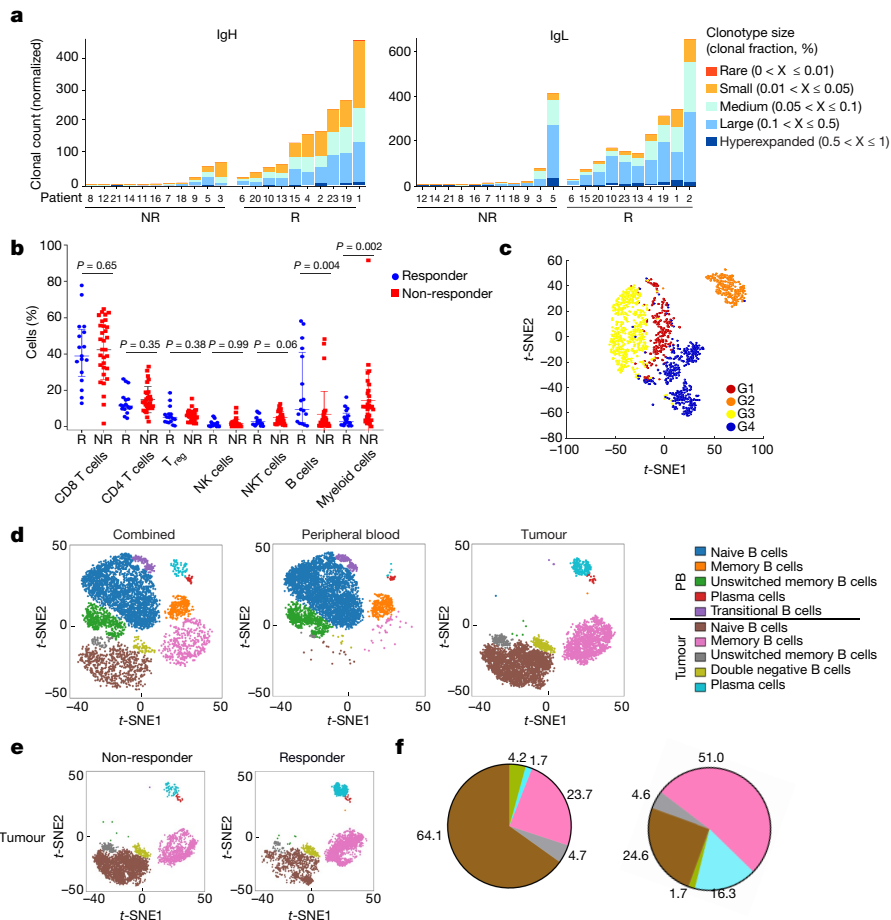
**Fig. 2 | TLSs containing B cells, T cells and follicular dendritic cells are predictive of response to ICB.** **a**, Quantification of CD20 cells by singlet immunohistochemistry and association with response to neoadjuvant ICB in resectable melanoma with responders defined as having complete or partial response by RECIST 1.1 and non-responders as having less than a partial response ( $n = 11$  NR and 10 R at baseline and  $n = 11$  NR and 9 R on treatment). **b**, **c**, Density of TLSs (**b**) and ratio of tumour area occupied by TLSs (**c**) and correlation by treatment response ( $n = 7$  NR and 7 R at baseline and  $n = 10$  NR and 8 R after treatment). For **a–c**, bars indicate median values, and errors bars

denote interquartile range; individual data points are shown. *P* values were determined by two-sided Mann–Whitney *U*-test. **d**, Representative image of CD20 staining of TLSs in a responder after treatment with ipilimumab and nivolumab. **e**, Additional staining of boxed area in **d** showing associated haematoxylin and eosin (H&E) staining and singlet immunostaining of CD20, CD8, CD4, FOXP3 and CD21. **f**, Multiplex immunofluorescence assay of TLSs as in **d** for the following markers: CD20, CD21, CD4, CD8, FOXP3 and DAPI. Original magnification,  $\times 20$ .

## BCR and single-cell RNA-seq offer functional insight

Next, we performed several in-depth analyses to gain insight into the phenotype and function of the infiltrating B cells, and how they might be contributing to responses to ICB. Reasoning that differences in the clonotypes of B cell receptors (BCRs) between responders and non-responders would be indicative of an anti-tumour B cell response, we probed our RNA-seq data for BCR sequences using the modified TRUST algorithm. In these studies, we identified significantly increased clonal counts for both immunoglobulin heavy and light chains (IgH and IgL;  $P = 0.001$  and  $P = 0.004$ , respectively) and increased BCR diversity in responders than in non-responders ( $P = 0.002$  and  $P = 0.0008$ ), which suggests an active role for B cells in anti-tumour immunity (Fig. 3a, Extended Data Fig. 8). To complement these analyses, we analysed single-cell RNA-seq data from baseline and on-treatment samples from an independent cohort of patients with metastatic melanoma treated with ICB ( $n = 48$  tumour samples; 1,760 B cells from 32 patients treated with PD1 blockade monotherapy, CTLA4 blockade monotherapy, or combined blockade of both PD1 and CTLA4, including samples from some patients in our neoadjuvant ICB-treated cohort<sup>44</sup>) (Supplementary Tables 22, 31). Similar to observations made in our clinical trial cohort, we found that B cells

were significantly enriched in tumours from responders versus non-responders and were predictive of a response (odds ratio 1.05,  $P = 0.02$ ) (Fig. 3b, Extended Data Fig. 9a, Supplementary Table 23). Unbiased analysis for markers of B cells (using all expressed genes in the CD45<sup>+</sup>CD19<sup>+</sup> population only) associated with clinical outcome demonstrated 46 markers were significantly enriched in lesions from responders and 147 markers were significantly enriched in non-responder lesions (Extended Data Fig. 9b, Supplementary Tables 24, 25). Pathways upregulated in responders as compared to non-responders include those consistent with increased immune activity such as CXCR4 signalling, cytokine receptor interaction and chemokine signalling pathways (Supplementary Table 26). Unsupervised clustering of B cells using *k*-means clustering, after testing for the robustness of each solution, identified four distinct B cell clusters, G1 (B cells, switched, activated IgD<sup>+</sup> cells), G2 (plasma cells), G3 (B cells unswitched IgD<sup>+</sup>) and G4 (B cells, switched, activated IgD<sup>+</sup> cells, with unique markers relative to G1), each of which is associated with different functional states (Fig. 3c, Extended Data Fig. 9c, Supplementary Tables 27, 28). No significant differences were identified when testing for associations of each individual cluster (G1–G4) with the clinical outcome, probably owing to limited sample size. Pathway analysis was also performed on bulk RNA-seq data from our clinical trial



**Fig. 3 | Analyses of B-cell receptor clones and single-cell analyses suggest active role for B cells in anti-tumour immunity.** **a**, Normalized clonal counts for BCRs identified in patients with high-risk resectable melanoma treated with neoadjuvant ICB. Both the IgH and IgL are evaluated with responders and non-responders as shown. All samples analysed at baseline. **b**, Scatter plots demonstrating the percentage of various cell types as indicated between responders ( $n = 17$ ) and non-responders ( $n = 31$ ) from a separate cohort of patients with advanced melanoma analysed by single-cell RNA-seq. Samples before and after treatment are combined. B cells are represented by the CD45<sup>+</sup>CD19<sup>+</sup> population. Data are median values with interquartile ranges, and individual data points are shown.  $P$  values were determined by two-sided Mann–Whitney  $U$ -test. Adjustments for multiple comparisons were not made. **c**,  $t$ -distributed stochastic neighbour embedding ( $t$ -SNE) plot of all B cells collected and analysed by single-cell RNA-seq in **b**. Cells are coloured based on

four clusters identified by  $k$ -means clustering (G1–G4). Number of cells analysed is 1,760 B cells from 48 tumours arising in 32 patients treated with PD1 blockade monotherapy, CTLA4 blockade monotherapy, or combined PD1 and CTLA4 blockade. **d**,  $t$ -SNE plots demonstrating peripheral blood and intratumoral combined B cell populations from mass cytometric analyses in responders versus non-responders ( $n = 4$  R and 4 NR for peripheral blood and  $n = 5$  R and 3 NR for tumour) from the neoadjuvant ICB trial in patients with advanced melanoma. **e**, Intratumoral B cell phenotypes included in **d** grouped by response. **f**, Quantification of B cell subtypes in **e**. Plots in **d–f** represent combined analyses of tumours ran simultaneously with the peripheral blood samples ( $n = 5$  R and 3 NR) and include baseline and on-treatment samples as described in Supplementary Table 31. Statistical analyses including all samples are presented in Extended Data Fig. 10b.

cohort, revealing increased immune signalling pathways in responders than in non-responders, including T cell receptor signalling, major histocompatibility complex-mediated antigen presentation and processing, differentiation of T helper 1 and 2 (T<sub>H</sub>1 and T<sub>H</sub>2) cells, and costimulatory signalling associated with T cell signalling (Supplementary Tables 29, 30).

### CytoF shows differential B cell phenotypes

To gain further insight into the potential functional role of B cells in the response to ICB, we performed mass cytometry (CyTOF) in evaluable tumour and peripheral blood samples (seven responders and three non-responders for tumour, and four responders and four non-responders for peripheral blood from our neoadjuvant ICB trial). Sample size was limited owing to the amount of tumour available given prioritization for other studies as well as tumour viability. These analyses included patients with nodal and non-nodal metastases (Extended Data Fig. 10a, Supplementary Tables 31, 32).

We first assessed differences between intratumoral B cells and those in the peripheral blood of patients. In these studies, unique clusters of CD45<sup>+</sup>CD19<sup>+</sup> (B cell) populations including naive (CD19<sup>+</sup>, CD27<sup>+</sup>, IgD<sup>+</sup>), transitional (CD19<sup>+</sup>, CD24<sup>+</sup>, CD38<sup>+</sup>, CD10<sup>+</sup>, CD27<sup>+</sup>, IgD<sup>+</sup>), unswitched and switched memory (CD19<sup>+</sup>, CD27<sup>+</sup>, IgD<sup>+</sup>), double-negative (CD19<sup>+</sup>, CD27<sup>+</sup>, IgD<sup>+</sup>), and plasma (-like) cell (CD19<sup>+</sup>, CD20<sup>+</sup>, CD22<sup>+</sup>, CD38<sup>+</sup>, CD27<sup>+</sup>) populations were found in peripheral blood and tumour samples, with distinct profiles in the tumour compared with peripheral blood samples (Fig. 3d, Extended Data Figs. 10a, b, 11a, b). Intratumoral B cells had reduced expression of CD21, CD23, CD79b and CXCR5, pointing to distinct functional and migratory profiles compared to similar B cell populations in the peripheral blood (Extended Data Fig. 11b). We next compared the phenotypes of B cells in tumours and peripheral blood from responders and non-responders to ICB treatment. Although B cell subsets (naive, memory and transitional B cells and plasma cells) in the peripheral blood had a similar distribution in responders and non-responders (Fig. 3d, Extended Data Fig. 10b), significant differences



were noted in the subsets of B cells in tumours (Fig. 3e, f, Extended Data Fig. 10b). Specifically, tumours from responders had a significantly higher frequency of memory B cells, whereas non-responders had a significantly higher frequency of naive B cells ( $P=0.033$  for naive and  $P=0.033$  for memory) (Fig. 3e, f, Extended Data Fig. 10b). Other notable differences included an increase in plasma cells in responders compared with non-responders; however, this did not reach significance and was largely driven by data from one patient ( $P=0.3$ ) (Fig. 3e, f, Extended Data Fig. 10b). More granular characterization of the intratumoral B cells reveals an increased percentage of CXCR3<sup>+</sup> switched memory B cells ( $P=0.0083$ ) in responders than in non-responders; we also note increased CD86<sup>+</sup> B cells ( $P=0.017$ ) and increased germinal-centre-like (CD19<sup>+</sup>, CD20<sup>++</sup>, CD38<sup>+</sup>, CD27<sup>+</sup>, IgD<sup>+</sup>, CD86<sup>+</sup>, CD95<sup>+</sup>) B cells ( $P=0.24$ ) in responders as compared to non-responders (Extended Data Figs. 10c, d, 11c). Increased proliferation of B cells suggestive of germinal centre formation and activity is observed within TLSs (Extended Data Fig. 7d).

## Summary

In summary, we present multiomic data that support a role for B cells within TLSs in the response to ICB in patients with metastatic melanoma and RCC. Although the distinct mechanisms through which B cells contribute are incompletely understood, our data suggest that the same properties of memory B cells and plasma cells desirable for acquired immune responses may also be contributing to an effective T cell response after ICB. Importantly, these B cells are probably acting together with other key immune constituents of the TLS by altering T cell activation and function as well as through other mechanisms. Memory B cells may be acting as antigen-presenting cells, driving the expansion of both memory and naive tumour-associated T cell responses. B cells can also secrete an array of cytokines (including TNF, IL-2, IL-6 and IFN $\gamma$ ), through which they activate and recruit other immune effector cells, including T cells. The observation of switched memory B cells (that can differentiate into plasma cells) in responders suggests that they could be potentially contributing to the anti-tumour response by producing antibodies against the tumours. Although we did not have adequate samples to study this in our cohort, it is an important line of investigation moving forward, and insights could lead to new therapeutic approaches to enhance responses to ICB. Together, findings in these cohorts are provocative and represent important advances in our insight into therapeutic responses to ICB. Further studies are needed in additional (and larger) cohorts across tumour types and stage of disease, as well as with therapeutic regimens. These types of studies along with pre-clinical models will help lend statistical power to the notion that B cells independently contribute to anti-tumour immune function in the context of ICB therapy, and also to better understand the mechanisms through which B cells and TLSs may favourably affect responses. Nonetheless, findings from these unique cohorts provide important insight into the role of B cells and TLSs in therapeutic responses to ICB, and are likely to stimulate further research in this area.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-019-1922-8>.

- Chen, P. L. et al. Analysis of immune signatures in longitudinal tumor samples yields insight into biomarkers of response and mechanisms of resistance to immune checkpoint blockade. *Cancer Discov.* **6**, 827–837 (2016).
- Taube, J. M. et al. Association of PD-1, PD-1 ligands, and other features of the tumor immune microenvironment with response to anti-PD-1 therapy. *Clin. Cancer Res.* **20**, 5064–5074 (2014).

- Cottrell, T. R. & Taube, J. M. PD-L1 and emerging biomarkers in immune checkpoint blockade therapy. *Cancer J.* **24**, 41–46 (2018).
- Yarchoan, M., Hopkins, A. & Jaffee, E. M. Tumor mutational burden and response rate to PD-1 inhibition. *N. Engl. J. Med.* **377**, 2500–2501 (2017).
- Ayers, M. et al. IFN- $\gamma$ -related mRNA profile predicts clinical response to PD-1 blockade. *N. Clin. Invest.* **127**, 2930–2940 (2017).
- Subudhi, S. K. et al. Clonal expansion of CD8 T cells in the systemic circulation precedes development of ipilimumab-induced toxicities. *Proc. Natl Acad. Sci. USA* **113**, 11919–11924 (2016).
- Jacquelot, N. et al. Predictors of responses to immune checkpoint blockade in advanced melanoma. *Nat. Commun.* **8**, 592 (2017).
- Gopalakrishnan, V. et al. Gut microbiome modulates response to anti-PD-1 immunotherapy in melanoma patients. *Science* **359**, 97–103 (2018).
- Matson, V. et al. The commensal microbiome is associated with anti-PD-1 efficacy in metastatic melanoma patients. *Science* **359**, 104–108 (2018).
- Routy, B. et al. Gut microbiome influences efficacy of PD-1-based immunotherapy against epithelial tumors. *Science* **359**, 91–97 (2018).
- Fridman, W. H., Zitvogel, L., Sautès-Fridman, C. & Kroemer, G. The immune contexture in cancer prognosis and treatment. *Nat. Rev. Clin. Oncol.* **14**, 717–734 (2017).
- Colbeck, E. J., Ager, A., Gallimore, A. & Jones, G. W. Tertiary lymphoid structures in cancer: drivers of antitumor immunity, immunosuppression, or bystander sentinels in disease? *Front. Immunol.* **8**, 1830 (2017).
- Dieu-Nosjean, M. C., Goc, J., Giraldo, N. A., Sautès-Fridman, C. & Fridman, W. H. Tertiary lymphoid structures in cancer and beyond. *Trends Immunol.* **35**, 571–580 (2014).
- Sarvaria, A., Madrigal, J. A. & Saudemont, A. B cell regulation in cancer and anti-tumor immunity. *Cell. Mol. Immunol.* **14**, 662–674 (2017).
- Tsou, P., Katayama, H., Ostrin, E. J. & Hanash, S. M. The emerging role of B cells in tumor immunity. *Cancer Res.* **76**, 5597–5601 (2016).
- Sautès-Fridman, C., Petitprez, F., Calderaro, J. & Fridman, W. H. Tertiary lymphoid structures in the era of cancer immunotherapy. *Nat. Rev. Cancer* **19**, 307–325 (2019).
- Amaria, R. N. et al. Neoadjuvant immune checkpoint blockade in high-risk resectable melanoma. *Nat. Med.* **24**, 1649–1654 (2018).
- Becht, E. et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol.* **17**, 218 (2016).
- Yuen, G. J., Demissie, E. & Pillai, S. B lymphocytes and cancer: a love-hate relationship. *Trends Cancer* **2**, 747–757 (2016).
- Chiaruttini, G. et al. B cells and the humoral response in melanoma: the overlooked players of the tumor microenvironment. *Oncol Immunology* **6**, e1294296 (2017).
- Erdag, G. et al. Immunotype and immunohistologic characteristics of tumor-infiltrating immune cells are associated with clinical outcome in metastatic melanoma. *Cancer Res.* **72**, 1070–1080 (2012).
- Iglesia, M. D. et al. Genomic analysis of immune cell infiltrates across 11 tumor types. *J. Natl. Cancer Inst.* **108**, (2016).
- Ladányi, A. et al. Prognostic impact of B-cell density in cutaneous melanoma. *Cancer Immunol. Immunother.* **60**, 1729–1738 (2011).
- Garg, K. et al. Tumor-associated B cells in cutaneous primary melanoma and improved clinical outcome. *Hum. Pathol.* **54**, 157–164 (2016).
- Ladányi, A. et al. Density of DC-LAMP<sup>+</sup> mature dendritic cells in combination with activated T lymphocytes infiltrating primary cutaneous melanoma is a strong independent prognostic factor. *Cancer Immunol. Immunother.* **56**, 1459–1469 (2007).
- Martinet, L. et al. High endothelial venules (HEVs) in human melanoma lesions: Major gateways for tumor-infiltrating lymphocytes. *Oncol Immunology* **1**, 829–839 (2012).
- Avram, G. et al. The density and type of MECA-79-positive high endothelial venules correlate with lymphocytic infiltration and tumour regression in primary cutaneous melanoma. *Histopathology* **63**, 852–861 (2013).
- Messina, J. L. et al. 12-Chemokine gene signature identifies lymph node-like structures in melanoma: potential for patient selection for immunotherapy? *Sci. Rep.* **2**, 765 (2012).
- Goc, J. et al. Dendritic cells in tumor-associated tertiary lymphoid structures signal a Th1 cytotoxic immune contexture and license the positive prognostic value of infiltrating CD8<sup>+</sup> T cells. *Cancer Res.* **74**, 705–715 (2014).
- Posch, F. et al. Maturation of tertiary lymphoid structures and recurrence of stage II and III colorectal cancer. *Oncol Immunology* **7**, e1378844 (2017).
- Siliņa, K. et al. Germinal centers determine the prognostic relevance of tertiary lymphoid structures and are impaired by corticosteroids in lung squamous cell carcinoma. *Cancer Res.* **78**, 1308–1320 (2018).
- Cipponi, A. et al. Neogenesis of lymphoid structures and antibody responses occur in human melanoma metastases. *Cancer Res.* **72**, 3997–4007 (2012).
- Selitsky, S. R. et al. Prognostic value of B cells in cutaneous melanoma. *Genome Med.* **11**, 36 (2019).
- Griss, J. et al. B cells sustain inflammation and predict response to immune checkpoint blockade in human melanoma. *Nat. Commun.* **10**, 4186 (2019).
- Blank, C. U. et al. Neoadjuvant versus adjuvant ipilimumab plus nivolumab in macroscopic stage III melanoma. *Nat. Med.* **24**, 1655–1661 (2018).
- Cabrera, R. L. et al. Tertiary lymphoid structures improve immunotherapy and survival in melanoma. *Nature* <https://doi.org/10.1038/s41586-019-1914-8> (2020).
- Petitprez, F. et al. B cells are associated with survival and immunotherapy response in sarcoma. *Nature* <https://doi.org/10.1038/s41586-019-1906-8> (2020).
- Amaria, R. N. et al. Neoadjuvant plus adjuvant dabrafenib and trametinib versus standard of care in patients with high-risk, surgically resectable melanoma: a single-centre, open-label, randomised, phase 2 trial. *Lancet Oncol.* **19**, 181–193 (2018).
- Reddy, S. M. et al. Poor response to neoadjuvant chemotherapy correlates with mast cell infiltration in inflammatory breast cancer. *Cancer Immunol. Res.* **7**, 1025–1035 (2019).
- Song, I. H. et al. Predictive value of tertiary lymphoid structures assessed by high endothelial venule counts in the neoadjuvant setting of triple-negative breast cancer. *Cancer Res. Treat.* **49**, 399–407 (2017).

41. Germain, C. et al. Presence of B cells in tertiary lymphoid structures is associated with a protective immunity in patients with lung cancer. *Am. J. Respir. Crit. Care Med.* **189**, 832–844 (2014).
42. Cancer Genome Atlas Network. Genomic classification of cutaneous melanoma. *Cell* **161**, 1681–1696 (2015).
43. Cancer Genome Atlas Research Network. Comprehensive molecular characterization of clear cell renal cell carcinoma. *Nature* **499**, 43–49 (2013).
44. Sade-Feldman, M. et al. Defining T cell states associated with response to checkpoint immunotherapy in melanoma. *Cell* **175**, 998–1013 (2018).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2020

<sup>1</sup>Department of Surgical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. <sup>2</sup>Department of Breast Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. <sup>3</sup>Department of Genitourinary Cancers, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. <sup>4</sup>Department of Genomic Medicine, The University of Texas MD Anderson Cancer Center, Houston, TX, USA.

<sup>5</sup>Department of Stem Cell Transplantation and Cellular Therapy, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. <sup>6</sup>Department of Medicine, Massachusetts General Hospital Cancer Center, Boston, MA, USA. <sup>7</sup>Broad Institute of the Massachusetts Institute of Technology, Boston, MA, USA. <sup>8</sup>Department of Immunology, The University of

Texas MD Anderson Cancer Center, Houston, TX, USA. <sup>9</sup>Department of Bioinformatics and Computational Biology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. <sup>10</sup>Department of Melanoma Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. <sup>11</sup>Department of Cancer Biology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. <sup>12</sup>Department of Thoracic / Head and Neck Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. <sup>13</sup>Parker Institute for Cancer Immunotherapy, San Francisco, CA, USA. <sup>14</sup>Department of Pathology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. <sup>15</sup>Department of Translational and Molecular Pathology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. <sup>16</sup>Immunology Program, Sloan Kettering Institute, Memorial Sloan Kettering Cancer Center, New York, NY, USA. <sup>17</sup>Nanostring Technologies, Seattle, WA, USA. <sup>18</sup>Division of Molecular Oncology and Immunology, The Netherlands Cancer Institute, Amsterdam, The Netherlands. <sup>19</sup>Departments of Medicine, Surgery, Immunology and Clinical and Translational Science, University of Pittsburgh, Pittsburgh, PA, USA. <sup>20</sup>Department of Epigenetics and Molecular Carcinogenesis, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. <sup>21</sup>INSERM, Cordeliers Research Center, Team Cancer, Immune Control and Escape, Paris, France. <sup>22</sup>University Paris Descartes Paris 5, Sorbonne Paris Cite, Centre de Recherche des Cordeliers, Paris, France. <sup>23</sup>Programme Cartes d'Identité des Tumeurs, Ligue Nationale Contre le Cancer, Paris, France. <sup>24</sup>These authors contributed equally: Beth A. Helmink, Sangeetha M. Reddy, Jianjun Gao, Shaojun Zhang, Rafet Basar. <sup>25</sup>These authors jointly supervised this work: Katayoun Rezvani, Padmanee Sharma, Michael T. Tetzlaff, Linghua Wang, Jennifer A. Wargo. \*e-mail: bhelmink@mdanderson.org; jwargo@mdanderson.org

### Patient cohort(s) and sample collection

For the melanoma neoadjuvant cohort (NCT02519322)<sup>17</sup>, 23 patients enrolled in a phase II clinical trial of neoadjuvant ICB. Twelve patients received nivolumab monotherapy with 3 mg kg<sup>-1</sup> every 2 weeks for up to 4 doses, and 11 patients received ipilimumab 3 mg kg<sup>-1</sup> with nivolumab 1 mg kg<sup>-1</sup> every 3 weeks for up to 3 doses followed by surgical resection. These patients were treated at the University of Texas MD Anderson Cancer Center and had tumour samples collected and analysed under Institutional Review Board (IRB)-approved protocols (2015-0041, 2012-0846). Of note, these studies were conducted in accordance with the Declaration of Helsinki and approved by the UT MD Anderson Cancer Center IRB. Response was defined as achieving a complete or partial radiographic response by RECIST 1.1 between pre-treatment imaging and post-neoadjuvant treatment imaging before surgical resection. Tumour samples were collected at several time-points for correlative studies including baseline and on-treatment (weeks 3 and 5 for nivolumab monotherapy, weeks 4 and 7 for combination ipilimumab with nivolumab). Tumour samples were obtained as core, punch or excisional biopsies performed by treating clinicians or an interventional radiologist. Samples were immediately formalin-fixed and paraffin-embedded (FFPE), snap-frozen or digested following tissue collection.

Additional patients off-protocol included five patients with widely metastatic melanoma who were treated at the University of Texas MD Anderson Cancer Center and had tumour samples collected and analysed under IRB-approved protocols (LAB00-063 and PA17-0261). Samples were immediately FFPE after tissue collection.

For the validation melanoma cohort, we used samples of 18 patients enrolled in the OpACIN-neo trial (NCT02437279). In the phase 1b OpACIN-neo trial, 20 patients with palpable stage III melanoma were randomized 1:1 to receive ipilimumab 3 mg kg<sup>-1</sup> and nivolumab 1 mg kg<sup>-1</sup>, either 4 courses after surgery (adjuvant arm), or 2 courses before surgery and two courses post-surgery (neoadjuvant arm). Coprimary endpoints were safety/feasibility and tumour-specific expansion of T cells. For this current correlative study, response was defined as not having disease relapse. These patients were treated at the Netherlands Cancer Institute (Amsterdam). The study was conducted in accordance with the Declaration of Helsinki and approved by the medical ethics committee of the Netherlands Cancer Institute. All subjects provided informed consent before their participation in the study. Patients underwent a pre-treatment tumour biopsy (1× formalin-fixed and paraffin-embedded (FFPE) and 2× fresh frozen) obtained as a core biopsy performed by a radiologist. RNA was extracted from one frozen biopsy for RNA-seq analysis. We included only 18 patients in our analysis because the tumour purity in the frozen pre-treatment biopsy of 2 patients was too low, therefore no RNA could be isolated and these patients could not be included in this analysis. The clinical responses of this cohort have been previously described<sup>35</sup>.

The RCC trial was an open-label, randomized, pre-surgical/pre-biopsy trial (NCT02210117) in which adults with metastatic RCC without previous immune checkpoint therapy and anti-VEGF therapy were enrolled and randomized 2:3:2 to receive nivolumab (3 mg kg<sup>-1</sup> once every 2 weeks, ×3 doses), nivolumab plus bevacizumab (3 mg kg<sup>-1</sup> once every 2 weeks ×3 plus 10 mg kg<sup>-1</sup> ×3) or nivolumab plus ipilimumab (3 mg kg<sup>-1</sup> once every 2 weeks ×3 plus 1 mg kg<sup>-1</sup> ×2), followed by surgery (cytoreductive nephrectomy or metastasectomy), or biopsy at week 8–10, and subsequent nivolumab maintenance therapy for up to 2 years. Response was assessed at 8 weeks and then at ≥12 weeks by RECIST 1.1 criteria. Clinical response data collection is still ongoing. For this current correlative study, clinical response for primary endpoint analysis was defined as achieving a complete or partial response at ≥12 weeks. Blood and tumours before and after treatment were obtained for correlative studies by IRB-approved laboratory protocol PA13-0291. Tumour samples were obtained as core biopsies or surgical resection performed by

interventional radiologists or surgeons. Samples were immediately FFPE or snap-frozen after tissue collection.

The single-cell RNA-seq B cell analysis used a dataset from 32 patients with metastatic melanoma ( $n = 48$  samples) treated with anti-PD1 ( $n = 37$ ), anti-CTLA4 ( $n = 2$ ), or anti-PD1 and anti-CTLA4 ( $n = 9$ )<sup>44</sup>. Patient response was determined by RECIST criteria: complete response and partial response for responders, or stable disease and progressive disease for non-responders. For the analysis, we focused on individual lesions and classified them into two categories: responder ( $n = 17$ ) including complete-response and partial-response samples; non-responder ( $n = 31$ ) including stable-disease and progressive-disease samples, based on radiological tumour evaluations. Samples were collected after patients provided a written consent for research and genomic profiling of collected tissue as approved by the Dana-Farber/Harvard Cancer Center Institutional Review Board (DF/HCC protocol 11-181) and UT MD Anderson Cancer Center (LAB00-063 and 2012-0846).

For the targeted therapy cohort, 13 patients received neoadjuvant and adjuvant dabrafenib and trametinib as part of a single-centre, open-label randomized phase 2 trial for patients with BRAF(V600E) or BRAF(V600K) (that is, Val600Glu or Val600Lys)-mutated melanoma (NCT02231775)—8 weeks of neoadjuvant oral dabrafenib 150 mg twice per day and oral trametinib 2 mg per day followed by surgery, then up to 44 weeks of adjuvant dabrafenib plus trametinib starting 1 week after surgery for a total of 52 weeks of treatment<sup>38</sup>. Patient radiographic response was determined by RECIST criteria with stable disease (non-responders) and partial response or complete response (responders) noted and coded as indicated; and pathological complete response determined by absence of residual viable malignant cells on H&E staining. These patients were treated at the University of Texas MD Anderson Cancer Center and had tumour samples collected and analysed under IRB-approved protocols. These studies were conducted in accordance with the Declaration of Helsinki.

The authors confirm for all studies involving human research participants we have complied with all relevant ethical regulations.

### Gene expression profiling and analysis: RNA extraction for neoadjuvant melanoma ICB-treated cohort

Total RNA was extracted from snap-frozen tumour specimens using the AllPrep DNA/RNA/miRNA Universal Kit (Qiagen) following assessment of tumour content by a pathologist, and macrodissection of tumour bed if required. RNA quality was assessed on an Agilent 2100 Bioanalyzer using the Agilent RNA 6000 Nano Chip with smear analysis to determine DV200 and original RNA concentration. On the basis of RNA quality, 40–80 ng of total RNA from each sample then underwent library preparation using the Illumina TruSeq RNA Access Library Prep kit according to the manufacturer's protocol. Barcoded libraries were pooled to produce final 10–12 plex pools before sequencing on an Illumina NextSeq sequencer using one high-output run per pool of 76-bp paired-end reads, generating 8 fastq files (4 lanes, paired reads) per sample.

### RNA-seq data processing and quality check

RNA-seq FASTQ files were first processed through FastQC (v.0.11.5)<sup>45</sup>, a quality control tool to evaluate the quality of sequencing reads at both the base and read levels. The reads that had ≥15 contiguous low-quality bases (phred score < 20) were removed from the FASTQ files. STAR 2-pass alignment (v.2.5.3)<sup>46</sup> was then performed on the filtered FASTQ files with default parameters to generate RNA-seq BAM file for each sequencing event. After that, RNA-SeQC (v.1.1.8)<sup>47</sup> was run on the aligned BAM files to generate a series of RNA-seq related quality control metrics including read counts, coverage, and correlation. A matrix of Spearman correlation coefficients was subsequently generated by RNA-SeQC among all sequencing events. The correlation matrix was carefully reviewed and the sequencing event generated from one library pool

that showed poor correlation with other library pools from the same RNA sample were removed before sample-level merging of BAM files.

### Gene expression quantification and normalization

HTSeq-count (v.0.9.1)<sup>48</sup> tool was applied to aligned RNA-seq BAM files to count for each gene how many aligned reads overlap with its exons. The raw read counts generated from HTSeq-count (v.0.9.1)<sup>48</sup> were normalized into fragments per kilobase of transcript per million mapped reads (FPKM) using the RNA-seq quantification approach suggested by the bioinformatics team of NCI Genomic Data Commons (GDC; <https://gdc.cancer.gov/about-data/data-harmonization-and-generation/genomic-data-harmonization/high-level-data-generation/rna-seq-quantification>). In brief, FPKM normalizes read count by dividing it by the gene length and the total number of reads mapped to protein-coding genes using a calculation described below:

$$\text{FPKM} = \frac{\text{RC}_g \times 10^9}{\text{RC}_{pc} \times L}$$

in which  $\text{RC}_g$  denotes the number of reads mapped to the gene;  $\text{RC}_{pc}$  denotes the number of reads mapped to all protein-coding genes; and  $L$  denotes the length of the gene in base pairs (calculated as the sum of all exons in a gene). The FPKM values were then  $\log_2$ -transformed for further downstream processes.

### RNA-seq analysis for OpACIN-neo trial

RNA-seq and data analysis were performed as previously described<sup>35</sup>.

### Affymetrix microarray for RCC

The Affymetrix microarray data were created using the Affymetrix Clariom D Assay (Human). There are 28 available pre-treatment samples from 3 arms: nivolumab ( $n = 6$ ), nivolumab plus bevacizumab ( $n = 14$ ) and nivolumab plus ipilimumab ( $n = 8$ ). The raw CEL files were normalized using the built-in SST-RMA method of the Affymetrix Transcriptome Analysis Console (TAC, v.4.0) software. The cell lineage scores were calculated using the R package MCP-counter algorithm (v.1.1.0). The Limma R software package<sup>49</sup> was used to identify DEGs from normalized microarray data for the RCC cohort.

### Identification of DEGs

The HTSeq normalized read count data for all expressed coding transcripts was processed by Deseq2 (v.3.6)<sup>50</sup> software to identify DEGs between two response (responders versus non-responders) groups. A cut-off of gene-expression fold change of  $\geq 2$  or  $\leq 0.5$  and a  $\text{FDR } q \leq 0.05$  was applied to select the most DEGs. The Limma R software package<sup>49</sup> was used to identify DEGs from normalized microarray data for the RCC cohort.

### Deconvolution of the cellular composition with MCP-counter

The R package software MCP-counter<sup>18</sup> was applied to the normalized  $\log_2$ -transformed FPKM expression matrix to produce the absolute abundance scores for eight major immune cell types (CD3<sup>+</sup> T cells, CD8<sup>+</sup> T cells, cytotoxic lymphocytes, natural killer cells, B lymphocytes, monocytic lineage cells, myeloid dendritic cells and neutrophils), endothelial cells and fibroblasts. The deconvolution profiles were then hierarchically clustered and compared across response and treatment groups.

### Pathway enrichment analyses

The network-based pathway enrichment analysis was performed using DEGs across responder and non-responder groups in the bulk-tissue RNA-seq data from the melanoma neoadjuvant cohort and single-cell RNA-seq data from the metastatic melanoma cohort. In the bulk-tissue, the differentially expressed genes that had a  $q < 0.05$  and  $\log_2$ -transformed fold change  $> 1.5$  or  $< -1.5$  were selected as input for network

based pathway enrichment analysis using ReactomeFiviz<sup>51</sup> application in Cytoscape<sup>52,53</sup>. In single-cell, the DEGs with  $q < 0.1$  were selected as input for pathway enrichment analysis. Pathway enrichment was calculated using several biological databases (KEGG, NCBI, Reactome, Biocarta and Panther) with hypergeometric test  $\text{FDR} < 0.01$ .

### TCGA SKCM and KIRC data downloading and patient selection

The normalized RNA-seq expression data of TCGA skin cutaneous melanoma (TCGA-SKCM) and Kidney Renal Clear Cell Carcinoma (TCGA-KIRC) was downloaded from NCI Genomic Data Commons (GDC; <https://portal.gdc.cancer.gov>) and the relevant clinical data were downloaded from recent TCGA PanCancer clinical data study<sup>54</sup>. The information of SKCM genomic subtypes was obtained from the TCGA-SKCM study<sup>42</sup>. To achieve a uniform cohort of patients with stage III (non-recurrent) melanoma for analysis, we applied an appropriate set of sequential filters: the TCGA-SKCM cohort was filtered to include patients with biospecimen tissue sites that included regional lymph node or regional subcutaneous metastases. We excluded patients presenting with stage IV disease. Then, to exclude patients with recurrent stage III disease, we excluded all patients for whom the number of days from the diagnosis of the primary to the accession date was more than 90 days. In addition, for a patient to be included, their tumour must also have had a defined melanoma driver type. Finally, we eliminated those lacking sufficient gene expression data, yielding a final stage III TCGA-SKCM cohort of  $n = 136$ . Survival data were missing for 9 of 136 samples, so  $n = 127$  samples were available for overall survival analyses. For TCGA-KIRC, the cases without available expression data were excluded and a total of 526 cases were taken into subsequent analysis.

### Survival analyses

In TCGA cohort, survival data were not available for nine samples and these were excluded from survival analysis. As previously described<sup>42</sup>, the survival time for each patient for the SKCM melanoma cohort was 'curated TCGA survival' (that is, from time of TCGA biospecimen procurement). The time to event was defined as the time interval from date of accession for each sample to date of death or censoring from any cause (curated value CURATED\_TCGA\_days\_to\_death\_or\_last\_follow-up; aka TCGA post-accession survival). The survival analysis was performed using Cox proportional hazards model and survival curves were plotted using Kaplan–Meier method. The statistical comparison of the survival curves was done using the log-rank test. The analysis was done using R package survival (<https://cran.r-project.org/web/packages/survival/index.html>).

### Statistical analyses

The statistical comparison between responder and non-responder groups for a given continuous variable was performed using two-sided Mann–Whitney  $U$ -test. The association between two continuous variables was assessed using Spearman's rank correlation coefficient. To control for multiple comparisons, we applied the Benjamini–Hochberg method<sup>55</sup> and calculated adjusted  $P$  values. Univariable and multivariable analysis predicting response to ICB was performed using logistic regression modelling. Biological replicates are indicated in the individual figure legends. Technical replicates were constrained to  $n = 1$  per time point, owing to limited tissue availability in patient-derived samples as well as prioritization for multiple studies. No statistical methods were used to predetermine sample size. The experiments were not randomized, and investigators were not blinded to allocation during experiments and outcome assessment unless stated otherwise.

### Single immunohistochemistry

H&E and immunohistochemistry staining were performed on FFPE tumour tissue sections. The tumour tissues were fixed in 10% formalin, embedded in paraffin, and serially sectioned. Four-micrometre sections were used for the histopathological study.

Sections were stained with mouse or rabbit anti-human monoclonal antibodies against CD20 (Dako, M0755, 1:1,400), CD21 (Novocastra, NCL-L-CD21-2G9, 1:10 or Leica, CD21-2G9; 1:20), CD23 (Leica, CD23-1B12, 1:15), CD4 (Novocastra, CD4-368-L-A, 1:80) CD8 (Thermo Scientific, MS-457-S, 1:25), FOXP3 (Biolegend, 320102, 1:50). All sections were counterstained with haematoxylin, dehydrated and mounted. All sections were processed with peroxidase-conjugated avidin/biotin and 3'-3'-diaminobenzidine (DAB) substrate (Leica Microsystem) and slides were scanned and digitalized using the scanscope system from Scanscope XT, Aperio/Leica Technologies.

Quantitative analysis of immunohistochemistry staining was conducted using the image analysis software ImageScope-Aperio/Leica. Five random areas (1 mm<sup>2</sup> each) were selected using a customized algorithm for each marker in order to determine the number of positive cells at high power field. The data are expressed as a density (total number of positive cells per mm<sup>2</sup> area). Immunohistochemistry staining was interpreted in conjunction with H&E stained sections.

## TLS quantification

TLSs were qualified and quantified using both H&E and CD20 immunohistochemistry staining. Structures were identified as aggregates of lymphocytes having histological features with analogous structures to that of lymphoid tissue with germinal centres (including B cells (CD19/20), T cells (CD3), follicular dendritic cells (CD21) and high endothelial venules (MECA79), appearing in the tumour area<sup>13,56–58</sup>. For the current study, criteria used for the quantification of TLS includes: (1) the total number of structures identified either within the tumoral area or in direct contact with the tumoral cells on the margin of the tumours (numbers of TLS per mm<sup>2</sup> area); and (2) a normalization of the total area occupied by the TLNs in relation of the total area of the tumour analysed (ratio: area of TLS/area tumour + TLNs).

## Multiplex immunofluorescence assay and analysis

For images shown in Fig. 2 and Extended Data Fig. 6, for immunofluorescence multiplex staining, we followed the staining method for the following markers: CD20 (Dako, M0755, 1:500) with subsequent visualization using fluorescein Cy3 (1:50); CD21 (Novocastra, NCL-L-CD21-2G9, 1:10) with subsequent visualization using fluorescein Cy5 (1:50); CD4 (CM153BK, Biocare, 1:25) with subsequent visualization using fluorescein Cy5.5 (1:50); CD8 (1:200, M7103, Dako) with subsequent visualization using fluorescein Cy3.5 (1:50); FOXP3 (Biolegend, 320102, 1:50) with subsequent visualization using fluorescein FITC (1:50) and nuclei visualized with DAPI (1:2,000). All of the sections were cover-slipped using Vectashield Hardset 895 mounting medium.

The slides were scanned using the Vectra slide scanner (PerkinElmer). For each marker, the mean fluorescent intensity per case was then determined as a base point from which positive calls could be established. For multispectral analysis, each of the individually stained sections was used to establish the spectral library of the fluorophores. Five random areas on each sample were analysed blindly by a pathologist at 20× magnification.

For additional multiplex images shown in Extended Data Fig. 5, for additional multiplex staining, we followed similar methods to the above for the following markers: MECA79-Dy550 (Novus, MECA-79, 1:100); CD20-Dy594 (Novus, IGEL/773; 1:100); CD4-AF647 (abcam, ERP6855, 1:100); and nuclei visualized with Syto13 at 500 nM. The slides were scanned with the GeoMx DSP machine as described below.

## GeoMx Digital Spatial Profiling: microscope and fluidics system overview

For immune profiling of T cells located within and outside TLS structures in patient samples, the GeoMx Digital Spatial Profiler (NanoString), a custom-built high-speed automated system and integrated instrument software, was used. A multiplexed cocktail of primary antibodies with UV photocleavable indexing oligonucleotides (GeoMx Immune Profile

Core; 22 targets, including 3 isotype controls and 4 additional modules; IO Drug Target, Immune Activation Status, Immune Cell Typing, and Pan Tumour) and 4 fluorescent markers was applied to a slide-mounted FFPE tissue section. For the fluorescent markers, we used Syto13 at 500 μM for nuclei visualization; CD20-Dy594 (Novus, IGEL/773; 1:100); CD3-AF647 (Novus, C3e/1308; 1:100); and PMEL-Dy550 (Novus, HMB45; 1:100) with S100B-Dy550 (Novus, 15F4NB; 1:100). Images at ×20 magnification were assembled to yield a high-resolution image of the tissue area of interest. The specific regions of interest (ROIs) for molecular profiling were then selected based on location (TLS or non-TLS areas of tumour) and CD3-positive staining and sequentially processed by the microscope automation. ROIs were selectively illuminated with UV light to release the indexing oligos by coupling UV LED light with a double digital mirror device (DDMD) module. Following each UV illumination cycle, the eluent was collected from the local region via microcapillary aspiration and transferred to an individual well of a microtiter plate. Once all ROIs were processed, pools of released indexing oligos were hybridized to NanoString optical barcodes for digital counting and subsequently analysed with an nCounter Analysis System.

## nCounter hybridization assay for photocleaved oligo counting

Hybridization of cleaved indexing oligonucleotides to fluorescent barcodes was performed using the nCounter Protein PlexSet reagents based on manufacturer's directions. Hybridizations were performed at 65 °C overnight in a thermocycler. After hybridization, samples were processed using the nCounter Prep Station and Digital Analyzer as per manufacturer instructions. Data were normalized to technical controls and area. Data were calculated against isotype controls to generate signal-to-noise ratios. Protein targets with a signal to noise ratio less than 2 were removed from downstream analysis.

## B cell clonotype analyses

The modified TRUST algorithm<sup>59</sup> was applied to extract the B cell immunoglobulin hypervariable regions from the bulk RNA-seq data and assemble the complementarity-determining region 3 (CDR3) sequences of the B cell heavy chain (IgH) and light chain (IgL). BCR clonotypes were identified and the clonal fraction was automatically calculated by TRUST. The output of TRUST was parsed by the R package tcR (v.3.4.1)<sup>60</sup> for downstream analyses. Only in-frame productive clonotypes were taken into subsequent analysis. The total number of BCR clonotypes detected per sample was normalized by the corresponding sequencing depth of each individual sample and calculated as per 100 million mapped reads. The top five clonotypes were selected by their clonal expression abundance. The BCR repertoire diversity was calculated by entropy from the tcR package<sup>60</sup>.

## Single-cell sequencing and analysis of CD45<sup>+</sup> B cells

Fresh isolated tumour samples were dissociated using the human tumour dissociation kit (Miltenyi Biotec; 130-095-929), sorted into 96-well plates containing 10 μl of TCL buffer (Qiagen) with 1% β-mercaptoethanol, using the following anti-human antibodies: FcX (Biolegend, 422302), CD45-PE (Biolegend, 304008), CD3-APC (Biolegend, 300412), CD235a-APC/Cy7 (Biolegend, 349116) and HLA-A,B,C-FITC (Biolegend, 311426). Sorting of viable cells was performed using the live/dead dye Zombie Violet (Biolegend, 77477). Single-cell libraries were generated using a modified version of the full-length Smart-seq2 protocol as previously described<sup>61</sup>, and were sequenced on a NextSeq 500 sequencer (Illumina), resulting in a median of approximately 1.4 million paired-end reads and a median of 2,588 genes detected per cell. A cutoff of log<sub>2</sub>(transcripts per million (TPM) + 1) ≥ 2 was used to define a gene as expressed in each single cell. For each sample, we computed the fraction of B cells using pre-defined markers (*CD19* and/or *MS4A1*). Notably, this is a plate-based protocol; thus, for each patient, we collected and sequenced the same number of cells (*n* = 384 CD45<sup>+</sup> cells per plate). Thus, the number of cells per patient



is equal, and the frequency reflects patients with either high or low B cell infiltrate.

### Unsupervised clustering of immune cells

To cluster all cells that passed quality control, we applied the  $k$ -means algorithm with a correlation distance metric, testing  $k = 3, \dots, 15$ . The algorithm was applied using all genes with variance  $> 6$ , yielding approximately 4,000 genes. This value was selected based on the relation between the variance and the fraction of cells expressing each gene. To determine the optimal number of clusters we applied the following steps: (1) we first examined how much of the complexity each cluster captures by applying the elbow method. This was done by computing the Pearson correlation matrix  $R$  and the distance matrix  $D$  as  $(1 - R)$ . We then computed the sum of pairwise distances between all cells in different clusters,  $\text{Dis}_b = \sum_{i=1}^k \left( \sum_{i \in C_i, j \notin C_i} D(i, j) \right)$ , and the total distance,  $\text{Dis}_t = \sum_{i,j} D(i, j)$ , in which  $i$  and  $j$  stand for each pair of single cells. The ratio between these two measures,  $V = \text{Dis}_b / \text{Dis}_t$  was used to estimate the variance explained by a given solution, such that in the extreme case in which all cells are clustered together or the case in which each cell is a single cluster, this ratio would be 0 and 1, respectively. Exploring this ratio, we then select the solutions that are near plateau ( $k = 10, \dots, 15$ ). (2) We then performed differential expression analysis (see 'Differential expression analysis') to search for gene markers that are significantly more highly expressed in a specific cluster as compared to all other clusters. Then, to avoid complex solutions, we excluded solutions with clusters that have too few marker genes ( $< 20$ ) distinguishing between them and the rest of the cells. (3) Finally, we performed a robustness analysis and selected the clustering solution with the highest median robustness score. Specifically, to determine the robustness of each clustering solution, we performed 100 iterations in which we randomly removed 10% of the cells, and re-ran the  $k$ -means algorithm and checked the stability of the clustering solution. We quantified the agreement of a given solution with the original one as the number of pairs of cells that were either clustered together, or not clustered together, in both solutions, divided by the total number pairs shared between the runs. This process yielded a median robustness measure of 0.96 for the selected  $k=11$ .

### Differential expression analysis

In all cases, differential expression analysis was applied to all genes that had an average expression level  $\log_2(\text{TPM} + 1) > 2$  in either tested groups,  $G_1$  and  $G_2$ . Then, for each gene  $i$ , we count the number of cells in  $G_1$  and  $G_2$  that express it with an expression level  $\log_2(\text{TPM} + 1) > 2$  or  $\leq 2$ . We then apply Fisher's exact test for the corresponding  $2 \times 2$  table. To identify significant differences, we considered genes with a Bonferroni-corrected  $q \leq 0.05$  and  $\log_2$ -transformed fold change  $> 0.5$ .

### CyTOF antibody conjugation

In-depth characterization of B cells from responders and non-responders was performed using metal-tagged antibodies. Metal conjugated antibodies were purchased from Fluidigm or conjugated to unlabelled antibodies in-house. All unlabelled antibodies were purchased in carrier-free form and conjugated with the corresponding metal tag using Maxpar X8 polymer per manufacturer's instructions (Fluidigm). Metal isotopes were acquired from Fluidigm and indium (III) chloride was acquired from Sigma-Aldrich. Antibody concentration was determined by measuring the amount of A280 protein using Nanodrop 2000 (Thermo Fisher Scientific). Conjugated antibodies were diluted using PBS-based antibody stabilizer supplemented with 0.05% sodium azide (Sigma-Aldrich) to a final concentration of  $0.5 \text{ mg ml}^{-1}$ . Antibodies used with the corresponding metal tag isotopes: CD45 (Fluidigm, H130,  $^{89}\text{Y}$ ), CD80 (Biolegend, 2D10,  $^{115}\text{In}$ ), CD138 (BD Biosciences, MI15,  $^{141}\text{Pr}$ ), CD19 (Fluidigm, HIB19,  $^{142}\text{Nd}$ ), CD5 (Fluidigm, UCHT2,  $^{143}\text{Nd}$ ), HLA-ABC (BD Biosciences, EMR8-5,  $^{144}\text{Nd}$ ), CD178 (Biolegend, NOK1,  $^{145}\text{Nd}$ ), IgD (Biolegend, IA6-2,  $^{146}\text{Nd}$ ), CD20 (Fluidigm, 2H7,  $^{147}\text{Sm}$ ), PDL1 (Fluidigm,

29E.2A3,  $^{148}\text{Nd}$ ), HLA-DR (Biolegend, L243,  $^{149}\text{Sm}$ ), CD25 (BD Biosciences, 2A3,  $^{150}\text{Nd}$ ), IGM (Biolegend, MHM-88,  $^{151}\text{Eu}$ ), CD95 (BD Biosciences, DX2,  $^{152}\text{Sm}$ ), CXCR5 (Fluidigm, RF8B2,  $^{153}\text{Eu}$ ), CD86 (BD Biosciences, IT2.2,  $^{154}\text{Sm}$ ), CD27 (Fluidigm, L128,  $^{155}\text{Gd}$ ), CXCR3 (Biolegend, G025H7,  $^{156}\text{Gd}$ ), CD10 (Fluidigm, H110a,  $^{158}\text{Gd}$ ), PDL-2 (Biolegend, 24F.10C12,  $^{159}\text{Tb}$ ), CD39 (Fluidigm, A1,  $^{160}\text{Gd}$ ), BAFF-R (Biolegend, 11C1,  $^{161}\text{Dy}$ ), CD79b (Fluidigm, CB3.1,  $^{162}\text{Dy}$ ), CD1d (Biolegend, 51.1,  $^{163}\text{Dy}$ ), CD23 (Fluidigm, EBVCS-5,  $^{164}\text{Dy}$ ), CD40 (Biolegend, 5C3,  $^{165}\text{Ho}$ ), CD24 (BD Biosciences, ML5,  $^{166}\text{Er}$ ), CD38 (BD Bioscience, HIT2,  $^{167}\text{Er}$ ), CD21 (Biolegend, Bu32,  $^{168}\text{Er}$ ), ICOS (Biolegend, C398.4A,  $^{169}\text{Tb}$ ), CTLA4 (Fluidigm, 14D3,  $^{170}\text{Er}$ ), CD9 (Biolegend, H19a,  $^{171}\text{Yb}$ ), CD11c (Biolegend, Bu15,  $^{172}\text{Yb}$ ), CD14 (Biolegend, HCD14,  $^{173}\text{Yb}$ ), PD1 (Miltenyi, PD1.3.1.3,  $^{174}\text{Yb}$ ), CXCR4 (Biolegend, 12G5,  $^{175}\text{Lu}$ ), CD22 (Biolegend, HIB22,  $^{176}\text{Yb}$ ), CD3 (Biolegend, UCHT-1,  $^{194}\text{Pt}$ ), cisplatin (Fluidigm,  $^{198}\text{Pt}$ ) and CD16 (Fluidigm, 3G8,  $^{209}\text{Bi}$ ).

### Sample preparation and acquisition

Peripheral blood mononuclear cells and tumour cells were collected and washed twice with wash buffer (0.5% bovine serum albumin (BSA) in PBS). For tumour, this included 9 responders and 9 non-responders, and for peripheral blood mononuclear cells, 8 responders and 8 non-responders. To determine the live population, cells were stained with  $1 \mu\text{M}$  cisplatin for 3 min. The reaction was stopped with FACS buffer (2% fetal bovine serum (FBS) in PBS), and the cells were washed once with wash buffer. Cells were then incubated with  $5 \mu\text{l}$  of Fc receptor blocking buffer reagent (Miltenyi) for 10 min at room temperature. Cells were incubated with surface antibodies at room temperature for 60 min, washed twice with wash buffer and stored overnight in 1 ml of 1.6% paraformaldehyde (EMD Biosciences) in PBS with 125 nM iridium nucleic acid intercalator (Fluidigm). The next day, samples were washed twice with cell staining buffer, re-suspended in 1 ml of MilliQ dH<sub>2</sub>O, filtered through a  $35\text{-}\mu\text{m}$  nylon mesh (cell strainer cap tubes, BD) and counted. Before analysis, samples were resuspended in MilliQ dH<sub>2</sub>O supplemented with EQ four element calibration beads at a concentration of  $0.5 \times 10^5$  per ml. Samples were acquired at 300 events per second on a Helios instrument (Fluidigm) using the Helios 6.5.358 acquisition software (Fluidigm).

### Data analysis

Mass cytometry data were normalized based on EQ four element signal shift over time using Fluidigm normalization software 2. Initial data processing was performed using Flowjo version 10.2. Mass cytometry data were normalized based on EQTM four element signal shift over time using Fluidigm normalization software 2. Initially, all responder and non-responder normalized FCS files were either concatenated or separately exported for downstream analyses. Data were processed and analysed using Cytobank; CD19<sup>+</sup> sample 'clean-up' was performed by gating on intact ( $^{91}\text{Ir}^+$  DNA stain), no beads ( $^{140}\text{Ce}^-$ ), live ( $^{198}\text{Pt}^-$ ), no T-cells CD3<sup>-</sup> ( $^{194}\text{Pt}$ ), no monocytes CD14<sup>-</sup> ( $^{173}\text{Yb}$ ) and CD45<sup>+</sup> ( $^{89}\text{Y}$ ), no natural killer cells CD16<sup>-</sup> ( $^{209}\text{Bi}$ ), CD19<sup>+</sup> B cells. Mass cytometry complex data were analysed using viSNE, in combination with heat map, to identify distinct subpopulations using the following parameters: CD19 ( $^{142}\text{Nd}$ ), CD20 ( $^{147}\text{Sm}$ ), CD5 ( $^{143}\text{Nd}$ ), HLA-ABC ( $^{144}\text{Nd}$ ), IgD ( $^{146}\text{Nd}$ ), PDL1 ( $^{148}\text{Nd}$ ), HLA-DR ( $^{149}\text{Sm}$ ), CD25 ( $^{150}\text{Nd}$ ), IgM ( $^{151}\text{Eu}$ ), CD95 ( $^{152}\text{Sm}$ ), CXCR5 ( $^{153}\text{Eu}$ ), CD86 ( $^{154}\text{Sm}$ ), CD27 ( $^{155}\text{Gd}$ ), CXCR3 ( $^{156}\text{Gd}$ ), CD10 ( $^{158}\text{Gd}$ ), CD39 ( $^{160}\text{Gd}$ ), BAFFR ( $^{161}\text{Dy}$ ), CD79b ( $^{162}\text{Dy}$ ), CD1d ( $^{163}\text{Dy}$ ), CD23 ( $^{164}\text{Dy}$ ), CD40 ( $^{165}\text{Ho}$ ), CD24 ( $^{166}\text{Er}$ ), CD38 ( $^{167}\text{Er}$ ), CD9 ( $^{171}\text{Yb}$ ), CD11c ( $^{172}\text{Yb}$ ), CXCR4 ( $^{173}\text{Lu}$ ), and CD22 ( $^{176}\text{Yb}$ ). Samples with fewer than 200 CD45<sup>+</sup>CD19<sup>+</sup> B cells were not used for downstream analyses. Percentages of different subpopulations of B cells were measured in aggregated responder and non-responder peripheral blood cells and tumour samples for each run; statistical analyses performed via unpaired Student's  $t$ -test.

### Analysis of peripheral blood exosomes from human plasma

Approximately 1 ml of plasma per patient sample contained in a cryovial was thawed rapidly in a  $37^\circ\text{C}$  water bath. The plasma was transferred

## Article

into a 1.5-ml Eppendorf tube and centrifuged at room temperature for 5 min at 800g and 10 min at 2,000g. The supernatant was filtered with a 0.22- $\mu$ m filter (6789-1302) directly into an ultracentrifuge tube (Z80615SCA, 331372). A distinct filter was used for each 500  $\mu$ l of plasma filtered, and each filter was subsequently cleared with  $2 \times 1$  ml PBS, all of which was collected into the ultracentrifuge tube. Additional PBS was added to the ultracentrifuge tube to reach 11 ml. The tubes were ultracentrifuged at 4 °C for 15–16 h at 100,000g using a Beckman Optima XE-90 ultracentrifuge. The pellet was resuspended in 200–300  $\mu$ l of PBS by pipetting up and down. The exosomes contained in this resuspension were stored at –80 °C until further use.

### Flow cytometric analyses of exosomes

Exosomes were thawed on ice. Concentration was determined using the NanoSight NS300 nanoparticle tracking analyser according to the manufacturer's directions, and 15  $\mu$ l of exosomes (which was equivalent to approximately  $4 \times 10^9$  particles on average) were mixed with 30  $\mu$ l of pre-washed anti-human CD63-coated Dynabeads (Invitrogen, 10606D). For one sample, the Nanosight measurement was erroneous and excluded. All samples were included in the flow cytometric analyses. Round-bottom 2-ml tubes were used. All pre-wash and washes thereafter were performed using 0.22  $\mu$ m filtered 0.1% BSA in PBS (0.1% BSA/PBS) and the samples were mixed well by pipetting up and down at each wash steps. One-hundred microlitres of 0.1% BSA/PBS was added to beads + exosomes mixture for a final volume of 145  $\mu$ l (15  $\mu$ l of exosomes + 30  $\mu$ l of Dynabeads + 100  $\mu$ l of 0.1% BSA/PBS). The samples were mixed by pipetting up and down and allowed to incubate for 4–16 h at room temperature on a benchtop rotator. Three-hundred microlitres of 0.1% BSA/PBS was added to the samples and the samples were placed on a magnet (1 min incubation minimum). The supernatant was discarded and the beads (and bound exosomes) were washed once with 400  $\mu$ l 0.1% BSA/PBS.

The beads (with bound exosomes) were resuspended in 400  $\mu$ l of 0.1% BSA/PBS and subsequently split into four distinct round-bottom 2-ml tubes, each containing 100  $\mu$ l. To each of these tubes, either antibodies or isotype control were added. These include: PE/Cy7 anti-human CD20 (Biolegend, 302312, clone 2H7) or isotype control PE/Cy7 mouse IgG2b (Biolegend, 400326, clone MCP-11); APC/Cy7 anti-human CD27 (Biolegend, 356424, clone M-T271) or isotype control APC/Cy7 mouse IgG1 (Biolegend, 400128, clone MOPC-21); PE/Cy7 anti-human CD9 (Biolegend, 312116, clone HI9a) or isotype control PE/Cy7 mouse IgG1 (Biolegend, 400126, clone MOPC-21); and Alexa Fluor 647 anti-human CD63 (Biolegend, 353016, clone H5C6) or isotype control Alexa Fluor 647 mouse IgG1 (Biolegend, 400130, clone MOPC-21). For each antibody or isotype control, 0.4  $\mu$ g was added to each tube. The samples were allowed to incubate at room temperature for 1–3 h, in the dark. Three-hundred microlitres of 0.1% BSA/PBS was added to the samples and the samples were placed on a magnet (1 min incubation). The supernatant was discarded and the beads (and bound exosomes) were washed once with 400  $\mu$ l 0.1% BSA/PBS. The beads were visible on the magnet at each step of the procedure described above. The supernatant was discarded and the beads were resuspended in 200  $\mu$ l of 0.1% BSA/PBS and transferred into flow cytometry tubes for flow cytometry analysis. The flow cytometry data were captured within 24 h of completing the staining of the beads-exosomes samples. If not read immediately after completing the staining, the flow cytometry tubes were stored at 4 °C in the dark. The data were subsequently analysed using FlowJo. Responder versus non-responder status was blinded until flow cytometry data capture and FlowJo analyses were completed.

For GPC1 staining, three tubes of beads with exosomes were processed in parallel. One tube did not receive any antibody (exosomes alone), one tube received primary antibody (1 h) followed by secondary antibody (1 h), and one tube received secondary antibody only (1 h). All three tubes were processed similarly, including for a wash step after one hour (post primary antibody incubation, 300  $\mu$ l 0.1% BSA/PBS was

added to the samples and the samples were placed on a magnet for 1 min incubation, and then resuspended into 100  $\mu$ l of 0.1% BSA/PBS), and again another hour later (after the secondary antibody incubation), before transferred into a FC tube. All incubations were carried out at RT and covered from light, and beads were visible at each step when placed on the magnet. Rabbit anti-human GPC1 antibody was used (Sigma, SAB2700282, 3  $\mu$ l per tube), and Alexa Fluor 488 conjugated goat anti-rabbit IgG (Invitrogen, A-11008, 2  $\mu$ l per tube) were used. The samples were analysed by flow cytometry.

### Nanoimager analyses

Beads with exosomes stained for flow cytometry analysis for CD63 (Alexa Fluor 647 anti-human CD63) or isotype control as described above (see 'Flow cytometric analyses of exosomes') were evaluated by using the Nanoimager S Mark I from ONI (Oxford Nanoimaging) with the lasers 405 nm/150 mW, 488 nm/200 mW, 561 nm/300 mW, 640 nm/1 W and dual emission channels split at 560 nm. Data were processed on NimOS (v.1.25) from ONI. In brief, 25  $\mu$ l of sample was spotted onto a slide (Fisher Scientific, 12-550-15), covered with a 1.5H coverslip (Zeiss, 474030-9000), and immediately placed on the stage. All images were captured using HILO mode (highly inclined and laminated optical sheet) at an illumination angle of 35.0° with a 10.0-ms exposure setting for 200 frames. To minimize photobleaching, the focal plane of the beads was found under the 405 nm laser at 37% power, then switched to the 640 nm laser at 25% power for image acquisition.

### Electron microscopy analyses

Bead only and beads with exosomes were prepared as described above ('Flow cytometric analyses of exosomes'). The samples were magnetized and resuspended in 50  $\mu$ l of 1% glutaraldehyde in PBS at 4 °C, or in 30  $\mu$ l of 0.1% BSA/PBS, and mixed with 30  $\mu$ l of warm (60 °C) 1% agarose in distilled water. The agarose-bead mixture was allowed to cool on ice, and the gels were cut into approximately 1-mm<sup>3</sup> pieces and placed in 1% glutaraldehyde in PBS at 4 °C. Fixed samples were washed in 0.1 M sodium cacodylate buffer and treated with 0.1% Millipore-filtered cacodylate buffered tannic acid, postfixed with 1% buffered osmium, and stained en bloc with 1% Millipore-filtered uranyl acetate. The samples were dehydrated in increasing concentrations of ethanol, infiltrated, and embedded in LX-112 medium. The samples were polymerized in a 60 °C oven for approximately 3 days. Ultrathin sections were cut in a Leica Ultracut microtome (Leica), stained with uranyl acetate and lead citrate in a Leica EM Stainer, and examined in a JEM1010 transmission electron microscope (JEOL) at an accelerating voltage of 80 kV. Digital images were obtained using AMT Imaging System (Advanced Microscopy Techniques). Two-sided Mann–Whitney *U*-test was used to determine significance.

### Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

### Data availability

The additional datasets generated during and/or analysed during the current study for Clinical Trial NCT02519322 are now available in the European Genome-phenome Archive repository (EGAS00001003178). Other datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

### Code availability

The authors declare that the code for reproducibility of data are publicly available or will be available upon request.

45. FastQC. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
46. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
47. DeLuca, D. S. et al. RNA-SeQC: RNA-seq metrics for quality control and process optimization. *Bioinformatics* **28**, 1530–1532 (2012).
48. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
49. Ritchie, M. E. et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
50. Love, M. I., Anders, S., Kim, V. & Huber, W. RNA-Seq workflow: gene-level exploratory analysis and differential expression. *F1000 Res.* **4**, 1070 (2015).
51. Wu, G., Dawson, E., Duong, A., Haw, R. & Stein, L. ReactomeFIViz: a Cytoscape app for pathway and network-based data analysis. *F1000 Res.* **3**, 146 (2014).
52. Smoot, M. E., Ono, K., Ruscheinski, J., Wang, P. L. & Ideker, T. Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* **27**, 431–432 (2011).
53. Shannon, P. et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
54. Liu, J. et al. An integrated TCGA Pan-cancer clinical data resource to drive high-quality survival outcome analytics. *Cell* **173**, 400–416 (2018).
55. Benjamini, Y., Drai, D., Elmer, G., Kafkafi, N. & Golani, I. Controlling the false discovery rate in behavior genetics research. *Behav. Brain Res.* **125**, 279–284 (2001).
56. Pimenta, E. M. & Barnes, B. J. Role of tertiary lymphoid structures (TLS) in anti-tumor immunity: potential tumor-induced cytokines/chemokines that regulate TLS formation in epithelial-derived cancers. *Cancers (Basel)* **6**, 969–997 (2014).
57. Hiraoka, N. et al. Intratumoral tertiary lymphoid organ is a favourable prognosticator in patients with pancreatic cancer. *Br. J. Cancer* **112**, 1782–1790 (2015).
58. Ruddle, N. H. Lymphatic vessels and tertiary lymphoid organs. *J. Clin. Invest.* **124**, 953–959 (2014).
59. Hu, X. et al. Landscape of B cell immunity and related immune evasion in human cancers. *Nat. Genet.* **51**, 560–567 (2019).
60. Nazarov, V. I. et al. tcR: an R package for T cell receptor repertoire advanced data analysis. *BMC Bioinformatics* **16**, 175 (2015).
61. Villani, A. C. et al. Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science* **356**, eaah4573 (2017).

**Acknowledgements** The correlative research was funded by philanthropic support from the MD Anderson Melanoma Moon Shot Program, the Parker Institute for Cancer Immunotherapy, the Dr. Miriam and Sheldon G. Adelson Medical Research Foundation, and the AIM at Melanoma Foundation. It was supported by The Immunotherapy Platform and the Core grant CA016672 (SMF) to support the Sequencing and Microarray Facility at UT-MDACC as well as the the Office of the Assistant Secretary for Defense for Health Affairs grant (W81XWH-16-1-0120 and W81XWH-16-1-0121). S.M.R. received support from National Institutes of Health T32 CA 009666 and Cancer Prevention Research Institute of Texas RR190020. B.A.H. received support from National Institutes of Health T32 CA 009599 and the MD Anderson Cancer Center support grant P30 CA016672. A.P.C. is supported by the CPRIT Research Training Program (RP170067) and the United States Department of State's Bureau of Educational and Cultural Affairs. E.Z.K. is supported by National Institutes of Health grant T32 CA009599. F.G.K. is supported with funding from the Odyssey Program at the MD Anderson Cancer Center with support from the Theodore N. Law Endowment for Scientific Achievement. Research in the Kalluri laboratory is supported by the Cancer Prevention and Research Institute of Texas and National Cancer Institute grants CA213233, P01CA117969, and CA195733. K.M.M. and M.A.Z. are supported by CPRIT RP190507. The High Resolution Electron Microscopy Facility at UT-MDACC is supported by CCSG grant NIGH P30CA016672. We thank to O. Contreras for technical support on the multiplex immunofluorescence and for M. Andrews for technical support on RNA-seq library preparation and K. McAndrews, D. P. Dowlatshahi, L. Snowden, J. Leveille and S. Yang for support with analyses of exosomes.

**Author contributions** B.A.H., S.M.R., J.G., S.Z. and R.B. contributed equally. Jointly supervising authors include P.S., K.R., M.T.T., L.W. and J.A.W. Corresponding authors are B.A.H. and J.A.W. J.A.W. designed the study. J.E.G., R.N.A., P.H., M.A.D., H.A.T., S.P., A.L., S. Woodman, C.U.B., T.N.S., P.S. and J.A.W. collected samples for correlative analyses. B.A.H., S.M.R., J.G., S.Z., R.B.,

R.T., K.Y., M.S.-F., J.B., G.H., V.G., Y.X., H.Z., A.P.C., W.L., V.S.L., F.G.K., A.L., R.A., S. Woodman, E.Z.K., P.-O.G., A.R., C.N.S., E.M.B., L.E.H., A.J.L., R.Z., C.W.H., D.A.L., S.O., M.B., S. Warren, D.R., O.K., E.A.R., D.P., C.U.B., T.N.S., L.H.B., M.A.Z., K.M.M., R.K., J.A., F.P., W.H.F., C.S.-F. and N.H. analysed and interpreted data. Specifically, K.Y., M.S.F. and N.H. assisted with single-cell RNA-seq analyses. S.O., M.B. and S. Warren assisted with Nanostring Digital Spatial Profiling analyses. R.B. and K.R. assisted with CyTOF. S.Z., R.T., L.W., M.A.Z. and K.M.M. assisted with BCR analyses. J.B., S.O., M.B., S. Warren, C.H., D.L. and M.T.T. assisted with singlet and multifocal immunohistochemistry. S.Z., L.W., J.E.G. and A.J.L. assisted with analyses of the TCGA patient cohorts. S.Z., R.T. and L.W. assisted with all bioinformatics and statistical analyses. R.T. performed pathway analyses. V.S.L., F.G.K. and R.K. assisted with analyses of exosomes. D.R., O.K., E.A.R., D.P., C.U.B. and T.N.S. provided assistance with analyses of patients from the OpACIN-neo trial. J.G., J.B. and P.S. assisted with analyses of the RCC patient cohort. All authors developed and approved the manuscript.

**Competing interests** J.A.W. is an inventor on a US patent application (PCT/US17/53717) submitted by the University of Texas MD Anderson Cancer Center that covers methods to enhance immune checkpoint blockade responses by modulating the microbiome. J.A.W. reports compensation for speaker's bureau and honoraria from Imedex, Dava Oncology, Omniplex, Illumina, Gilead, PeerView, Physician Education Resource, MedImmune and Bristol-Myers Squibb. J.A.W. serves as a consultant/advisory board member for Roche/Genentech, Novartis, AstraZeneca, GlaxoSmithKline, Bristol-Myers Squibb, Merck, Biothera Pharmaceuticals and Microbiome DX. J.A.W. also receives research support from GlaxoSmithKline, Roche/Genentech, Bristol-Myers Squibb, and Novartis. J.A.W., S.M.R. and B.A.H. are co-inventors on an unpublished patent application related to methods of targeting B cells to enhance response to immune checkpoint blockade. M.T.T. reports advisory board participation and speaker paid honorarium from Nanostring and Myriad Genetics. M.A.D. serves as a paid consultant for BMS, Novartis, and Roche/Genentech. M.A.D. also reports to be a principal investigator of a research grant from Roche/Genentech and an unpaid consultant to Nanostring. C.U.B. reports an advisory role in BMS, MSD, Roche, Novartis, GSK, AZ, Pfizer, GenMab and Pierre Fabre. C.U.B. receives research funding from BMS, Novartis and Nanostring. C.U.B. reports stock ownership from Uniti Cars and Neon Therapeutics. N.H. is a founder, stockholder and SAB member of Neon Therapeutics. W.H.F. serves as a consultant for AstraZeneca, Ipsen, Adaptimmune, OxfordBiotherapeutics, and Catalym. W.H.F. reports participation in data transparency committee for Servier and data management committee for Novartis. O.K. receives grant support from BMS. J.E.G. is a contributor of UpToDate-melanoma staging and prognosis. J.E.G. reports to be an unpaid member of Melanoma Research Foundation and Melanoma Research Alliance. J.E.G. reports to be on advisory board of Merck. R.K. reports to be scientific founder, stockholder and consultant of Codiak Biosciences. A.J.L. reports consultancies and research support from BMS, Genentech/Roche and MedImmune/ Astra-Zeneca. P.S. reports a patent licensed to Jounce Therapeutics. P.S. serves as a consultant for Constellation, Jounce Therapeutics, Neon, BioAtla, Pieris, Oncolytics, Forty-Seven, Polaris, Apricity, Marker, Codiak, ImaginAb, Hummingbird, Dragonfly, Lytx and Bristol-Myers Squibb (BMS). P.S. has ownership interests in Jounce Therapeutics, Neon, Constellation, Oncolytics, BioAtla, Forty-Seven, Apricity, Polaris, Marker, Codiak, ImaginAb, Hummingbird, Dragonfly and Lytx. L.H.B. reports to be on StemImmune/Calidi Scientific and Medical Advisory, Scientific Advisory Board of BoardNextCure, Replimmune, Western Oncolytics, Torque Therapeutics Khloris, Pyxis, Cytomix. L.H.B. reports to be Chair of Food and Drug Administration Cellular, Tissues and Gene Therapies Advisory Committee. R.Z. reports a patent application related to work on G1TR, PD1 and CTLA4. R.Z. is a consultant for Leap Therapeutics. P.H. is on advisory board for Dragonfly, GlaxoSmithKline, Immatics and Sanofi. S.W., S.O. and M.B. are employees and stockholders of NanoString Technologies. All other authors report no competing interests directly relevant to this work.

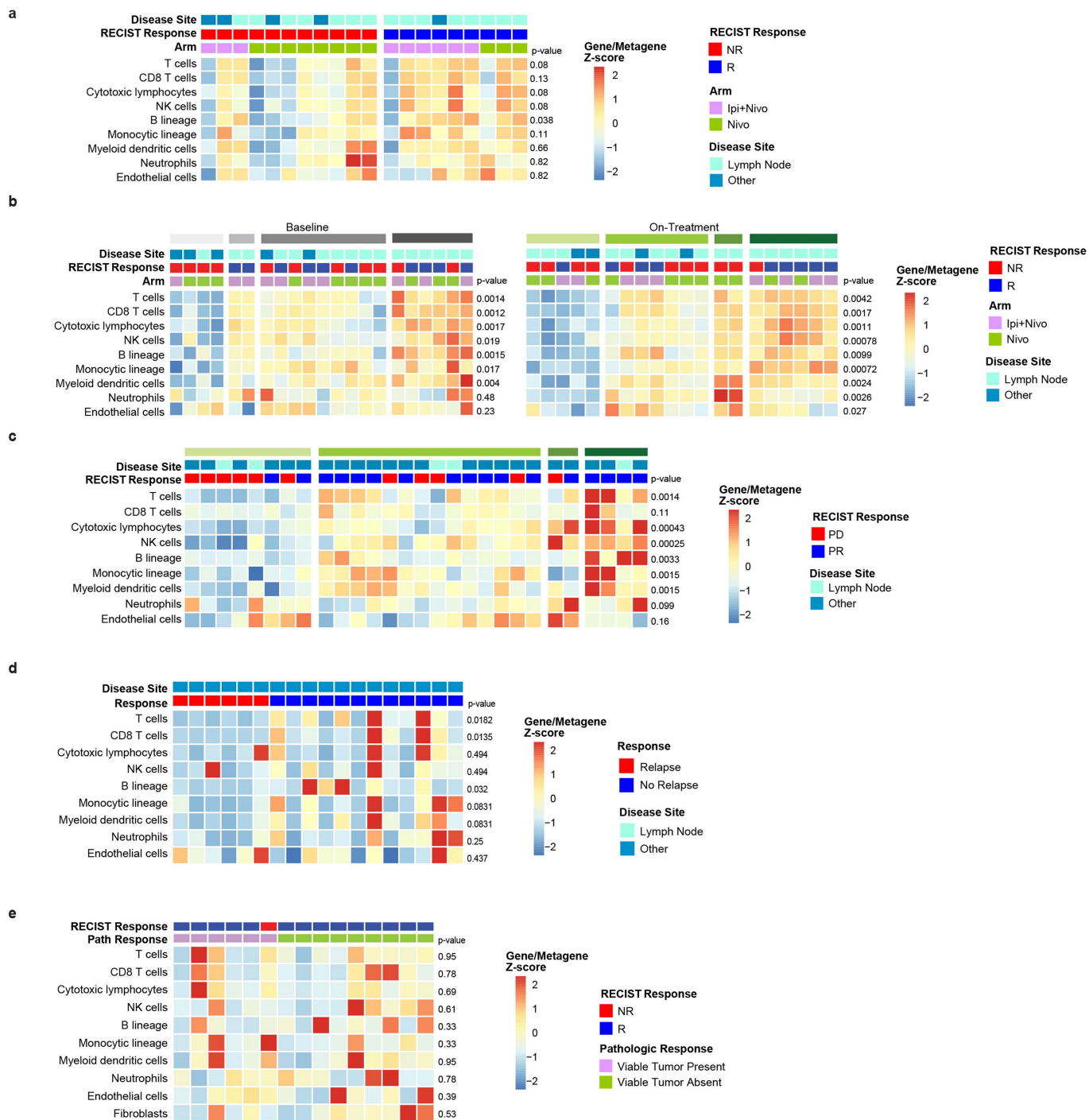
#### Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41586-019-1922-8>.

**Correspondence and requests for materials** should be addressed to B.A.H. or J.A.W.

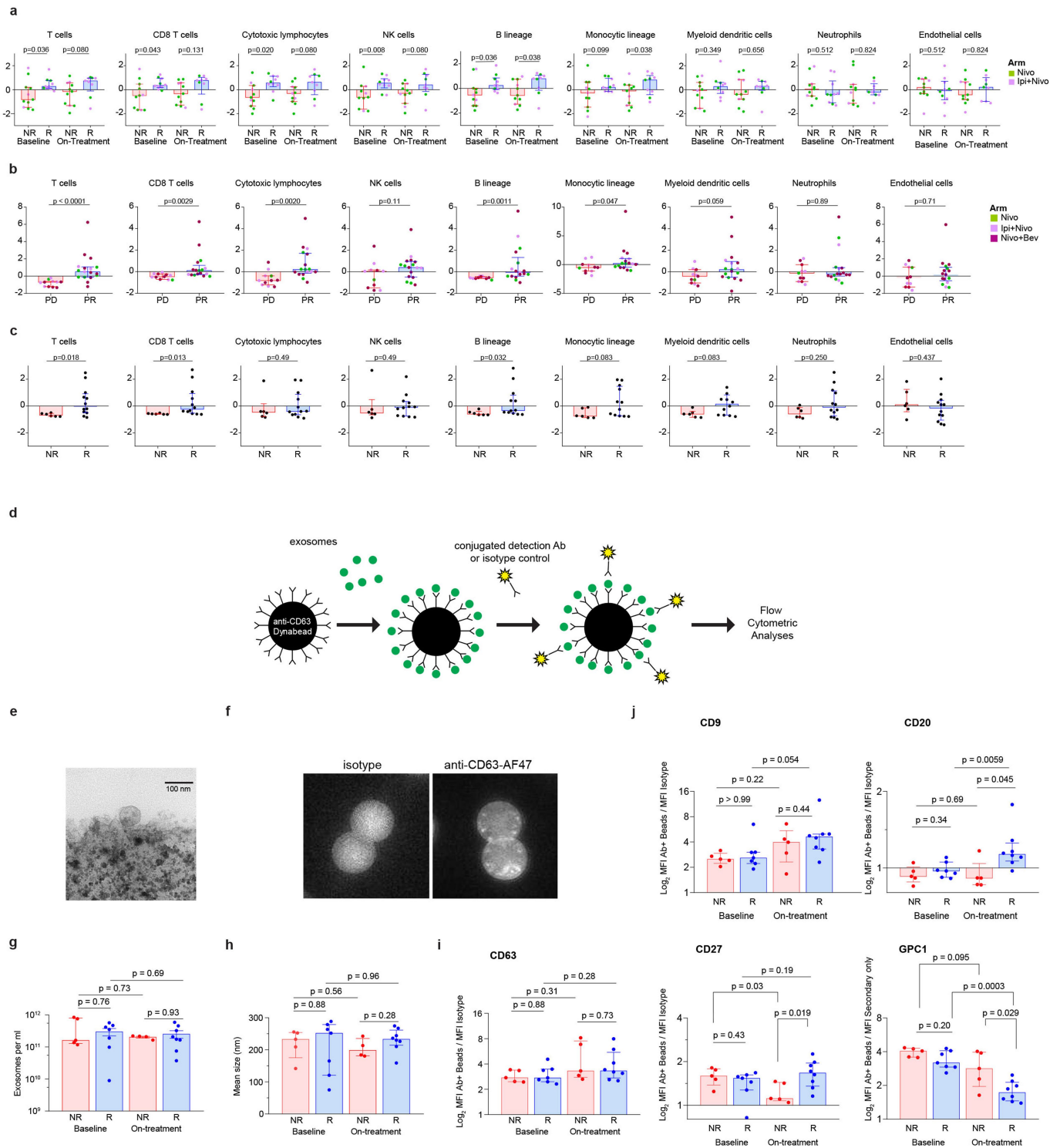
**Peer review information** *Nature* thanks James Mulé, Caroline Robert and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>.



**Extended Data Fig. 1 | MCP-counter results in patients with melanoma and RCC treated with pre-surgical ICB or targeted therapy. a**, Supervised clustering by response of MCP-counter scores in on-treatment samples from a cohort of high-risk patients with resectable melanoma treated with neoadjuvant ICB, with responders defined as achieving a complete or partial response by RECIST 1.1 ( $n = 11$  NR and 9 R). **b**, Analysis shown by unsupervised hierarchical clustering of baseline ( $n = 11$  NR and 10 R) and on-treatment samples ( $n = 11$  NR and 9 R) from the neoadjuvant melanoma cohort. Unique clusters identified are indicated by shaded boxes on top row. **c**, Unsupervised hierarchical analysis shown for metastatic RCC patients (same cohort as Fig. 1d;  $n = 11$  PD and 17 PR). Response (PR, partial response) or non-response (PD, progressive disease) as measured by RECIST 1.1. Unique clusters identified are indicated by shaded boxes on top row. **d**, Supervised clustering by response of

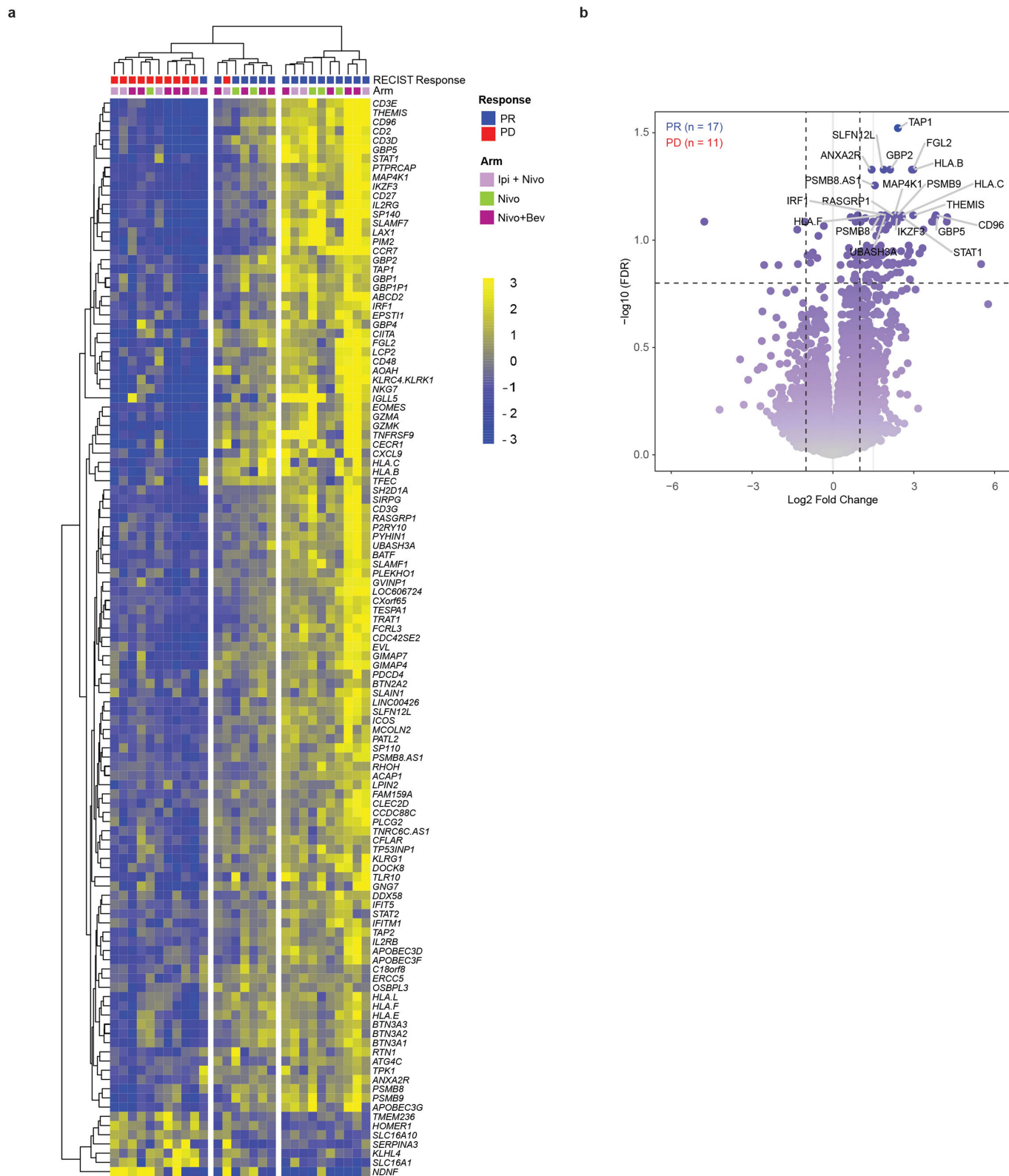
MCP-counter scores from OpACIN-neo clinical trial (NCT02437279) of neoadjuvant versus adjuvant ICB in high-risk resectable melanoma ( $n = 6$  NR and 12 R). Responders were defined as patients who did not have a relapse. **e**, Supervised clustering by response of MCP-counter scores in combined pre-treatment and on-treatment biopsies from a cohort of high-risk resectable melanoma patients treated with neoadjuvant targeted therapy (dabrafenib and trametinib) as part of NCT02231775 ( $n = 7$  patients for baseline and  $n = 8$  patients for on-treatment samples) with responder defined as achieving a complete or partial response by RECIST 1.1 and non-responder defined as having stable or progressive disease. Pathological response is defined by the presence or absence of viable tumour at time of surgical resection. *P* values were made using two-sided Mann–Whitney *U*-test.



**Extended Data Fig. 2 | Representation of MCP-counter scores for all patient cohorts and analyses of peripheral blood exosomes. a–c,** Box plot representation of heat maps for patients with: high-risk resectable melanoma treated with neoadjuvant ICB ( $n = 11$  NR and 10 R for baseline and  $n = 11$  NR and 9 R on treatment) as presented in Fig. 1c and Extended Data Fig. 1a, b (a); metastatic RCC treated with pre-surgical ICB as presented in Fig. 1d and Extended Data Fig. 1c ( $n = 11$  PD and 17 PR) (b); and high-risk resectable melanoma treated with ICB as part of OpACIN-neo trial as presented in Extended Data Fig. 1d ( $n = 6$  NR and 12 R) (c). For a–c, medians with interquartile range are shown. *P* values were determined by two-sided Mann–Whitney *U*-test. **d,** Schematic for exosomal analyses of serum samples from patients with melanoma on neoadjuvant ICB trial. **e,** Representative transmission electron

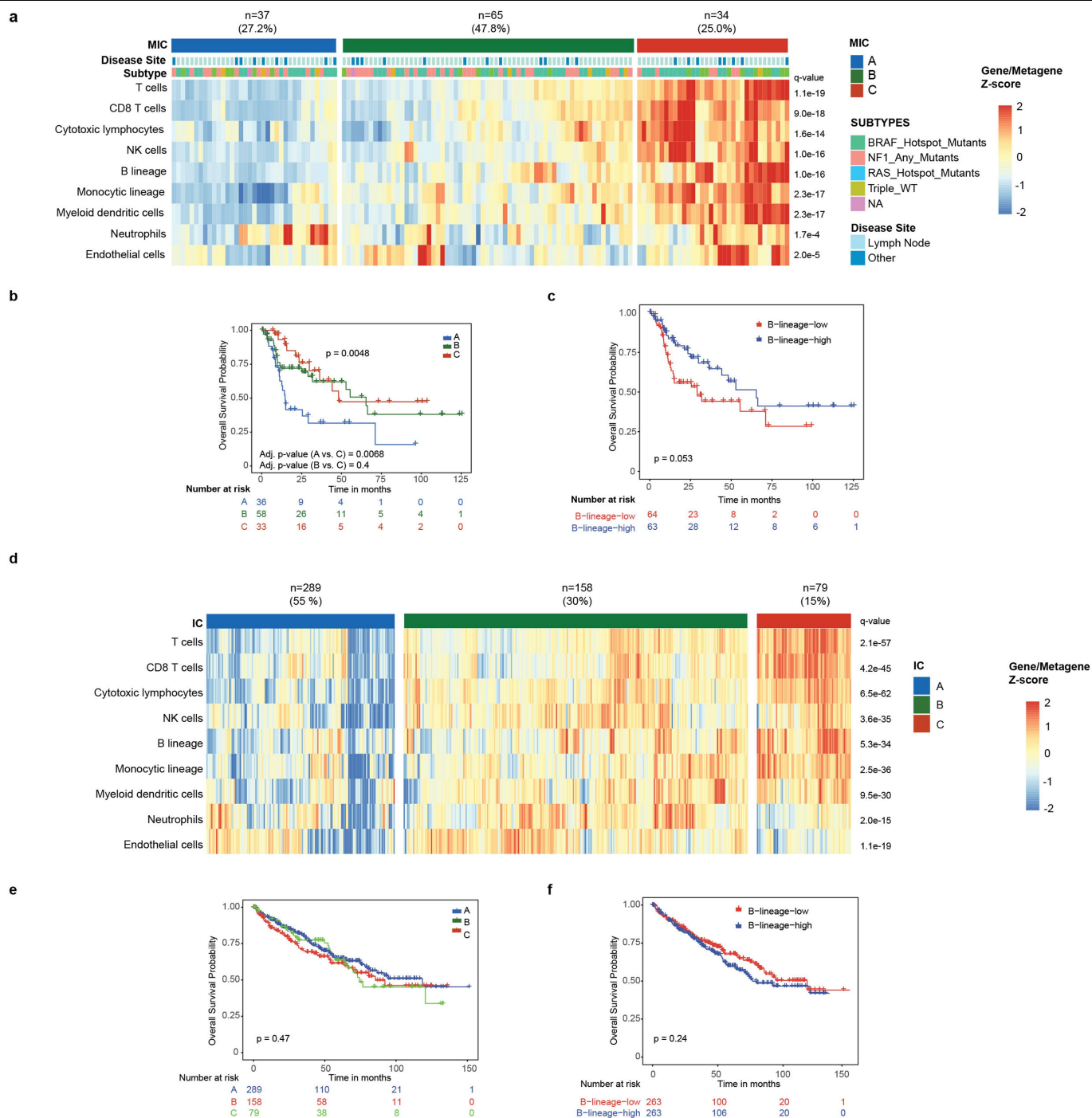
micrographs showing Dynabead with exosomes present after immunocapture. **f,** Nanoimager-captured images of the beads coated with CD63<sup>+</sup> exosomes as compared with isotype control. **g, h,** Exosomal concentration (g) and mean exosomal size (h) for serum samples for responders and non-responders at the time point indicated. **i,** Ratio of mean fluorescent intensity (MFI) of beads stained with anti-CD63 as compared to isotype control. **j,** Ratio of mean fluorescent intensity of beads stained with anti-CD9, -CD20, -CD27 and -GPC1 antibodies as compared to isotype control (or secondary antibody only for GPC1). For e–j, bars indicate median values and individual data points representing 8 R and 5 NR (unless indicated in the Methods) in addition to interquartile ranges. *P* values were determined using two-sided Mann–Whitney *U*-test.





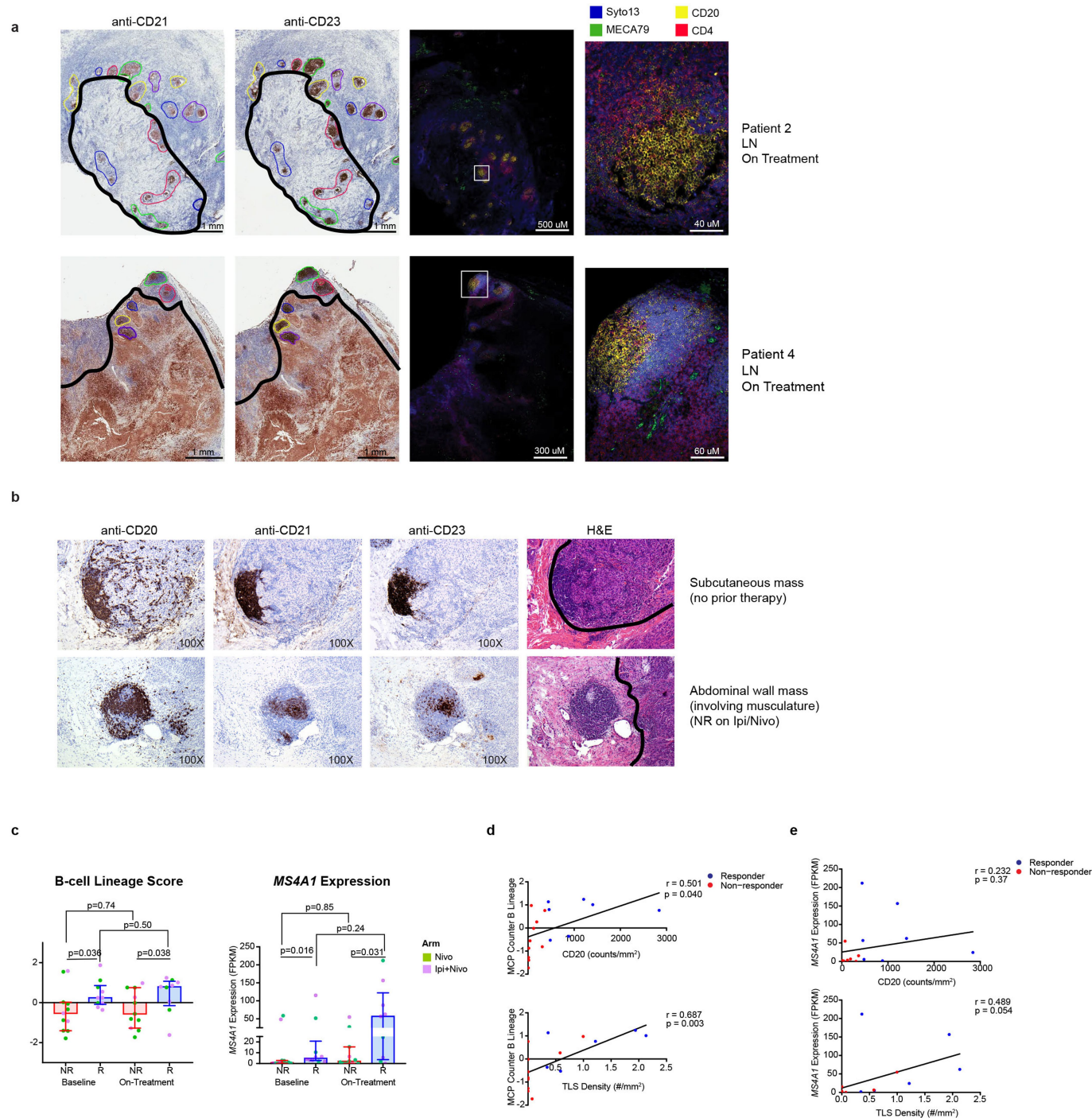
**Extended Data Fig. 3 | Transcriptional analysis of tumour specimens from patients with metastatic RCC treated with pre-surgical ICB. a,** Supervised hierarchical clustering by response of RCC tumour specimens at baseline of most DEGs by microarray analysis, with response defined as having a partial response by RECIST 1.1 and non-response as having progressive disease

(n = 11 PD and 17 PR). Fold change and P values are calculated by the limma package as described in the Methods. A cut-off of gene expression fold change of  $\geq 2$  or  $\leq 0.5$  and a FDR  $q \leq 0.05$  was applied to select DEGs. **b,** Volcano plot depiction of DEGs by response from same cohort.



**Extended Data Fig. 4 | Immune infiltrate is prognostic of improved disease-specific survival in TCGA cutaneous melanoma cohort but not the clear-cell RCC cohort.** **a**, Unsupervised hierarchical analysis of TCGA SKCM RNA-seq data using MCP-counter scores identifies three MICs with differential presence of individual cell types as indicated. Numbers of patients in each class is shown on top of the plot. *P* value determined by two-sided Kruskal–Wallis rank-sum test and *q* value calculated by FDR. **b**, Kaplan–Meier estimates of overall survival of MIC groups. **c**, Kaplan–Meier estimates of overall survival by B cell lineage scores shown by high and low groups dichotomized by median values. Overall survival was defined as the time interval from date of accession for each sample to date of death or censoring from any cause (Methods).

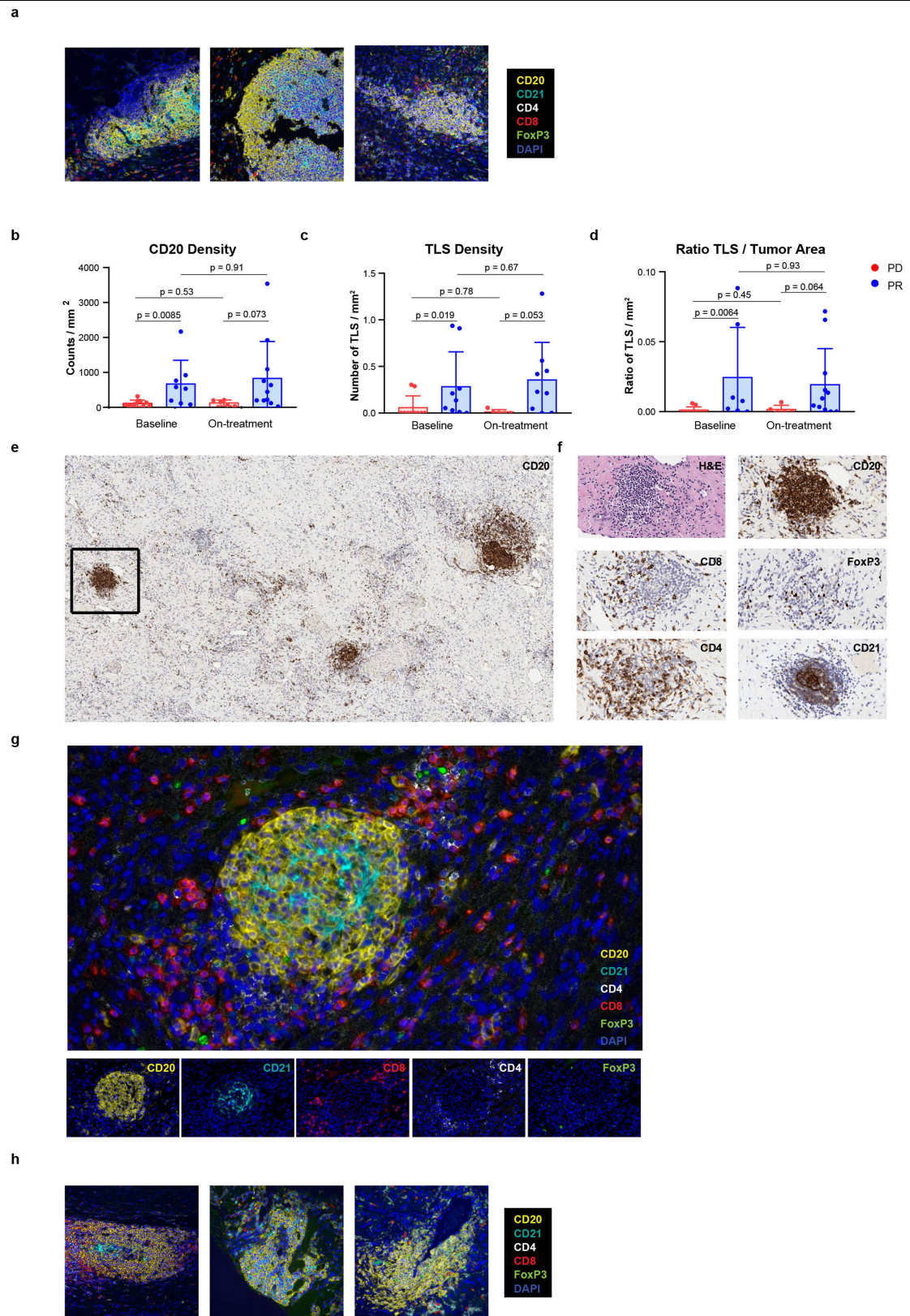
**d**, Unsupervised hierarchical analysis of TCGA KIRC RNA-seq data using MCP-counter scores identifies three immune classes with differential presence of individual cell types as indicated. Numbers of patients in each class are shown at top of plot. *P* values determined by two-sided Kruskal–Wallis rank-sum test *q* value calculated by FDR. **e**, Kaplan–Meier estimates of overall survival probability of immune class groups. **f**, Kaplan–Meier estimates of overall survival probability by B cell lineage scores shown by high and low groups dichotomized by median values. For both, overall survival was defined as the time interval from date of accession for each sample to date of death or censoring from any cause. For **b**, **c**, **e**, **f**, patient numbers are included in the table below the graph and *P* values were calculated by log-rank test.



**Extended Data Fig. 5 | TLSs found in nodal and non-nodal metastases are consistent with mature secondary follicular-like TLSs with modest correlation with gene expression data. a**, Representative TLSs in tumours from patients with melanoma treated with neoadjuvant ICB demonstrating maturation status as indicated by the presence of follicular dendritic cells (CD21) and germinal centre B cells (CD23). We also include multiplex immunohistochemistry for SYTO13, MECA79, CD20 and CD4 (with magnified view of individual TLSs indicated by white box on the right). Circles denote defined TLSs based on multiplex immunohistochemistry. Black line approximates tumour border. **b**, Representative TLSs from non-lymph node metastases on additional patients with metastatic melanoma indicated by H&E staining, as well as singlet staining for CD20, CD21 and CD23. Black line on H&E image denotes tumour border. **c**, Comparison of baseline and on-treatment

gene expression with MCP-counter analyses for B cell lineage as well as *MS4A1* expression by RNA-seq for patients with high-risk resectable melanoma treated with ICB as part of clinical trial ( $n = 11$  NR and 10 R for baseline and  $n = 11$  NR and 9 R for on-treatment). Response and treatment arm as indicated. Medians with interquartile range are shown. *P* values were determined by two-sided Mann-Whitney *U*-test. **d**, Linear regression modelling of MCP-counter values for B cell lineage with regards to CD20 counts ( $n = 10$  NR and 7 R) and TLS density ( $n = 10$  NR and 6 R) as indicated. **e**, Linear regression modelling of *MS4A1* gene expression with regards to CD20 counts ( $n = 10$  NR and 7 R) and TLS density ( $n = 10$  NR and 6 R) as indicated. These represent on-treatment time points. For **d**, **e**, *r*, values calculated by linear regression and *P* values for non-zero slope as calculated by Prism v.8.0.0.



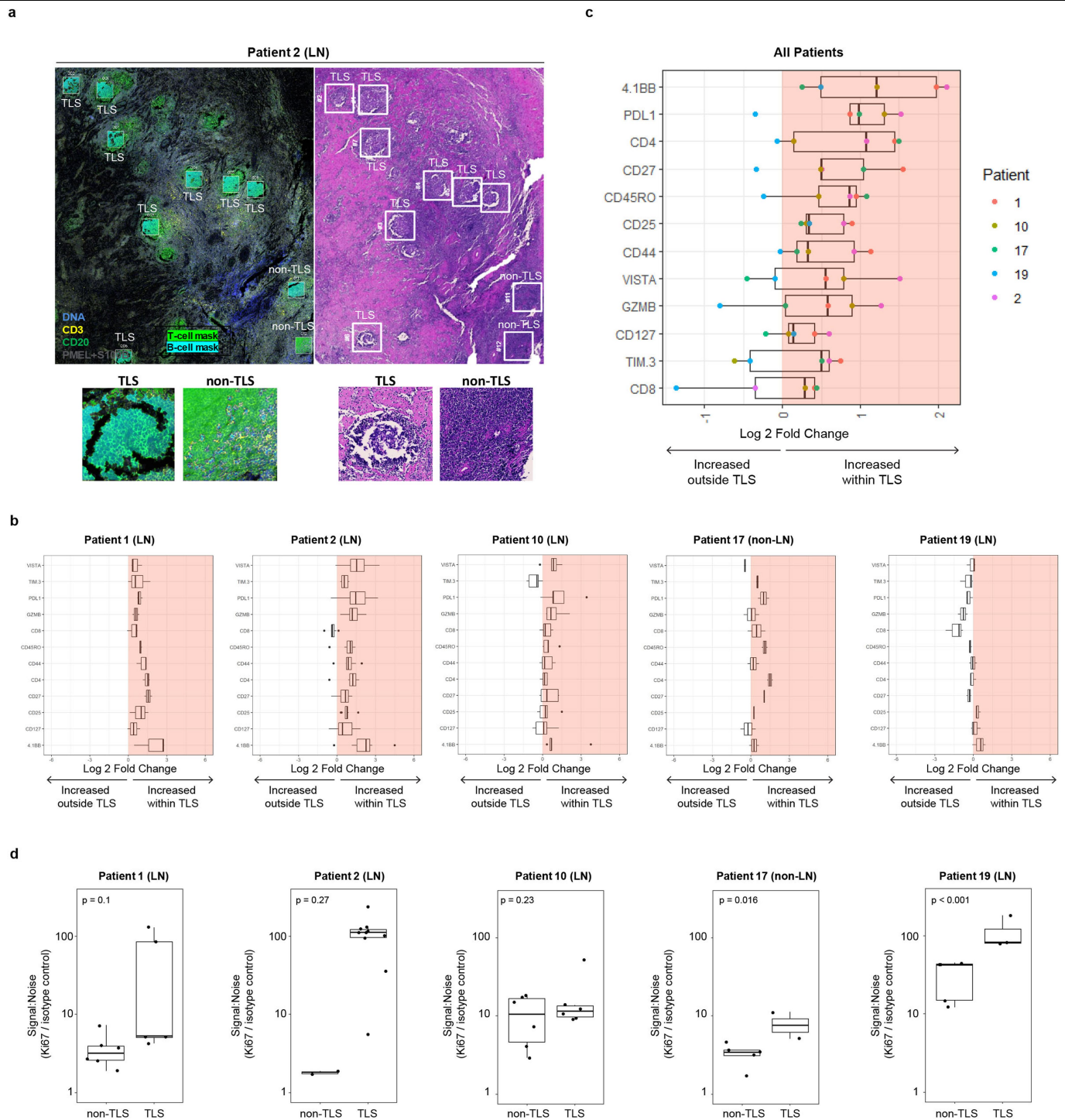


**Extended Data Fig. 6** | See next page for caption.

**Extended Data Fig. 6 | TLSs are associated with response in RCC similar to those observed in melanoma.** **a**, Multiplex immunohistochemistry images from three additional patients with advanced melanoma treated with neoadjuvant ICB. Staining as indicated and similar to Fig. 2. **b**, Quantification of CD20 cells by singlet immunohistochemistry and association with response to neoadjuvant ICB in metastatic RCC, with responders defined as having partial response and non-responders as having progressive disease by RECIST 1.1 ( $n = 10$  PD and 8 PR at baseline and  $n = 5$  PD and 11 PR on treatment). **c**, **d**, Density of TLSs ( $n = 10$  PD and 9 PR at baseline and  $n = 5$  PD and 9 PR on treatment) (**c**)

and ratio of tumour area occupied by TLSs ( $n = 10$  PD and 7 PR at baseline and  $n = 5$  PD and 11 PR on treatment) (**d**) and correlation by treatment response. Bars indicate median values and interquartile ranges are shown.  $P$  values were determined by two-sided Mann–Whitney  $U$ -test. **e–g**, Representative image of CD20 staining in responder with TLSs, associated H&E slide, singlet stains and characterization by multiplex immunofluorescence of TLSs. **h**, Multiplex immunohistochemistry images from three additional patients with RCC treated with pre-surgical ICB. Staining as indicated and similar to **g**.

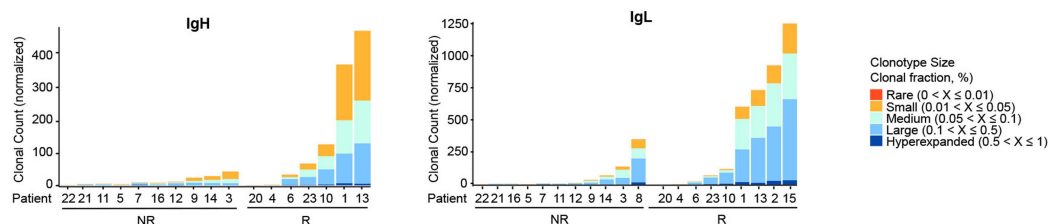




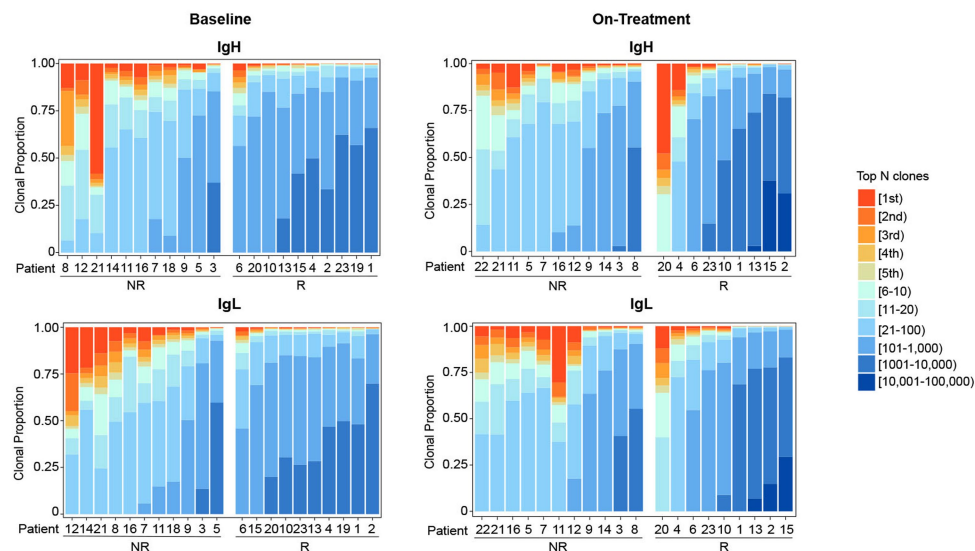
**Extended Data Fig. 7 | TLSs are associated with markers of T cell activation and response and B cell proliferation.** NanoString GeoMx Digital Spatial Profiling was used to perform high-plex proteomic analysis with spatial resolution. **a**, Example of selection of ROIs (200  $\mu\text{m} \times 200 \mu\text{m}$ ) from representative patients with melanoma treated with neoadjuvant ICB including ROI containing TLSs and ROIs outside the context of a TLS (non-TLS). ROI selection was completed using H&E staining and confirmed with immunofluorescence as shown using S100B and PMEL, SYTO13, CD3 and CD20. Masking for B cells and T cells as indicated based on CD3 and CD20 staining. **b**, Fold change ( $\log_2$ -transformation) in expression of various markers of T cell activation and response in TLS-associated T cells as compared to T cells found outside the TLS per individual slide. Data show individual TLS ROI values divided by the average non-TLS value of that slide. Increased expression in the context of TLSs is represented by shaded pink box ( $>0$ ). **c**, Average  $\log_2$ -

transformed fold change of expression for TLS-associated T cells as compared to non-associated T cells. Individual dots represent individual patients/slides as indicated. Data show the average  $\log_2$ -transformed count per TLS ROI value minus the average  $\log_2$ -transformed count per non-TLS ROI value per slide for each protein queried. For **b** and **c**, increased expression in the context of TLSs is represented by shaded pink box ( $>0$ ). Median and interquartile range are indicated. Error bars indicate 95% confidence intervals. **d**, Levels of Ki67 protein expression in B cell masks of non-TLSs and TLS ROIs by individual patient as indicated. Counts are represented as signal-to-noise ratios of Ki67 compared to geometric means of isotype controls. Median and interquartile range are indicated. Error bars indicated 95% confidence ratios, and  $P$  values were determined by Student's  $t$ -test. For **a-d**, the number of ROIs analysed for each patient are 11 for patient 1, 12 for patient 2, 12 for patient 10, 7 for patient 17 and 7 for patient 19.

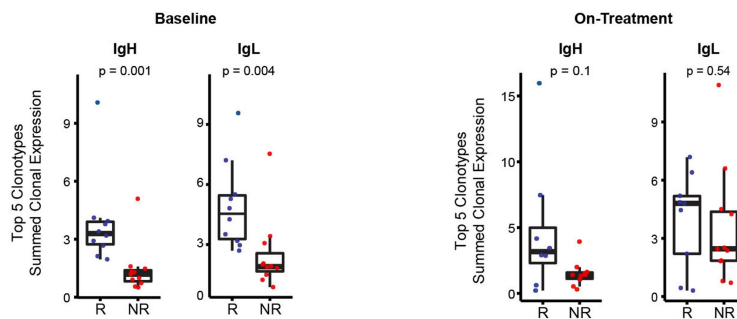
a



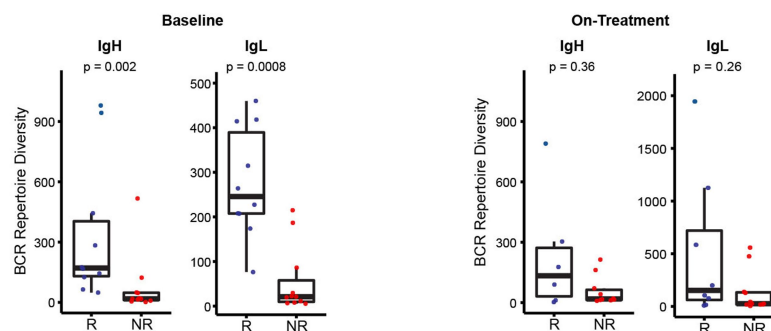
b



c



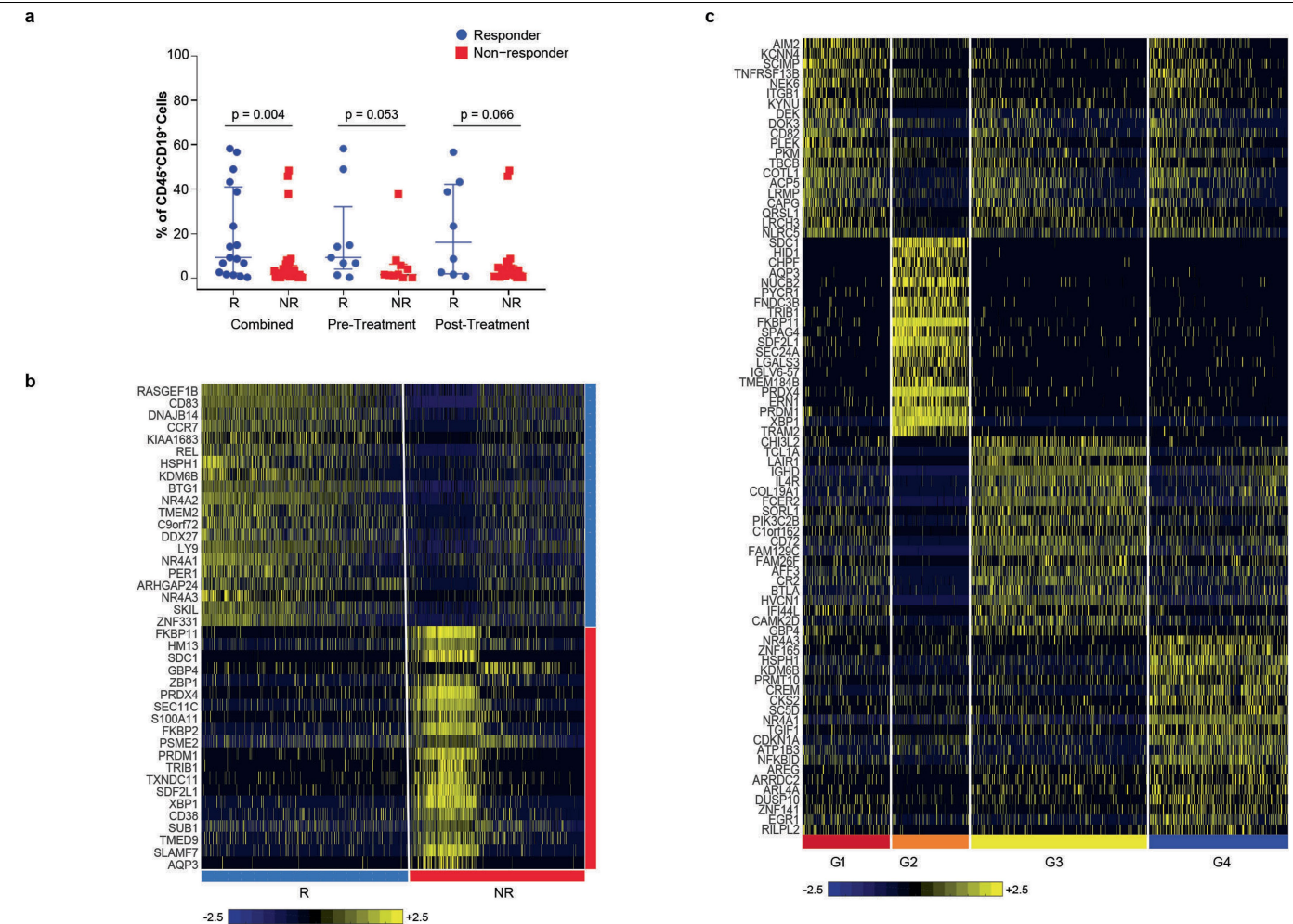
d



**Extended Data Fig. 8 | BCR analyses of intratumoral B cells in patients with advanced melanoma before and after treatment with neoadjuvant ICB.**

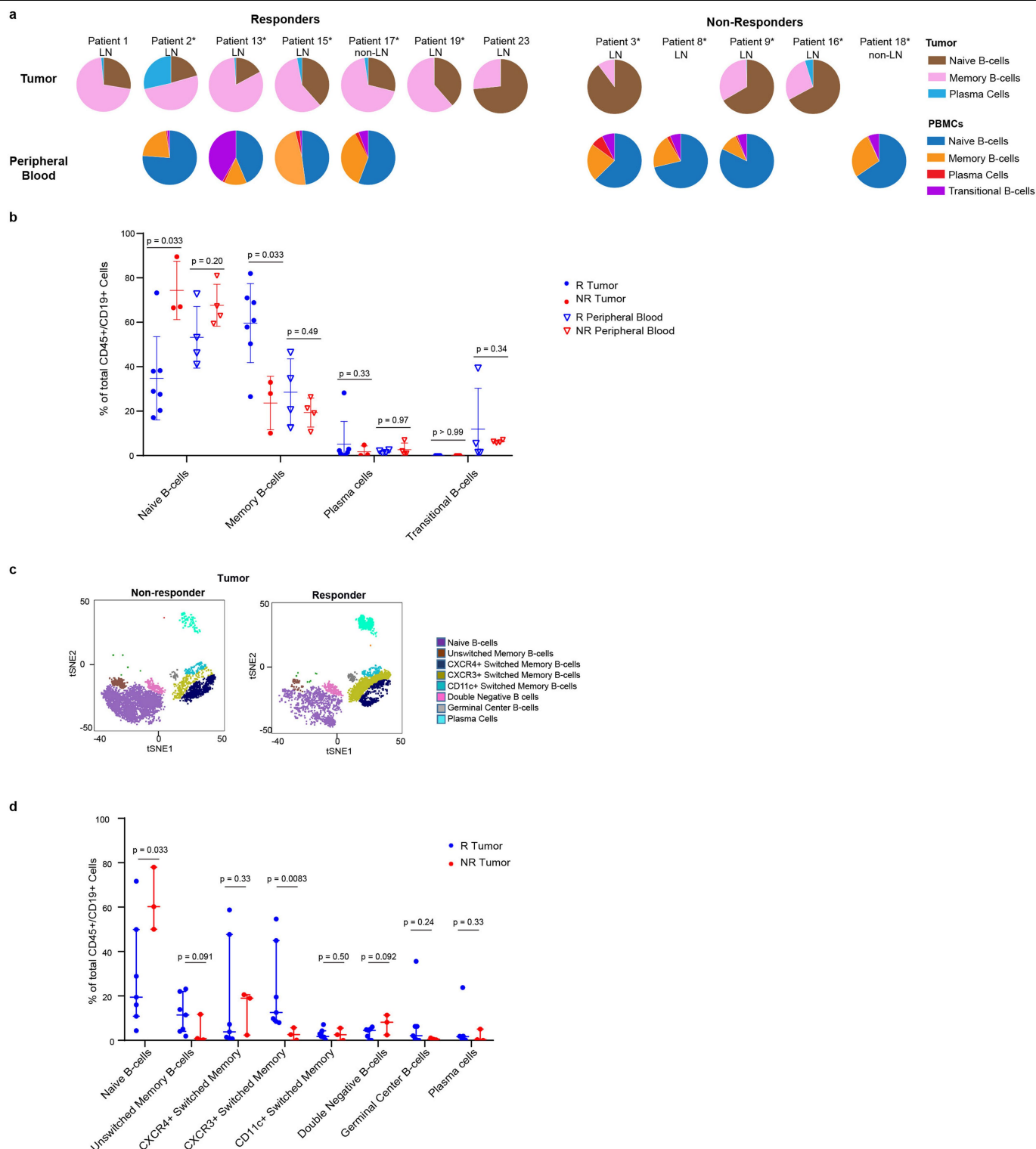
**a**, Clonal counts for all identified clonotypes for both IgH and IgL after treatment with ICB. Patients are grouped as responders and non-responders and identified as indicated in which each bar represents individual patient. **b**, Clonal proportion for all identified clonotypes for both IgH and IgL for baseline samples further evaluated in Fig. 3a and on-treatment samples in **a**.

Patients are grouped as responders and non-responders and identified as indicated in which each bar represents individual patient. **c**, **d**, Summed expression of top five clonotypes in normalized read counts (**c**) and BCR repertoire diversity (**d**) for responders and non-responders for both IgH and IgL at baseline ( $n = 11$  NR and 10 R for IgH and IgL) and on-treatment ( $n = 10$  NR and 9 R for IgH and  $n = 11$  NR and 9 R for IgL). Box plot shows median and interquartile range.  $P$  values determined by two-sided Mann-Whitney  $U$ -test.



**Extended Data Fig. 9 | Single-cell RNA-seq analysis reveals unique clusters of B cells associated with response to ICB. a,** Scatter plots comparing the percentage of CD45<sup>+</sup> cells staining positive for CD19<sup>+</sup> (B cells) as indicated between responder ( $n=17$ ) and non-responder ( $n=31$ ) samples with all time points combined or stratified by pre- and post-treatment as indicated. Data are median and interquartile range.  $P$  values were determined by two-sided Mann-Whitney  $U$ -test. **b,** Heat map displaying scaled expression values ( $\log_2(\text{TPM} + 1)$ )

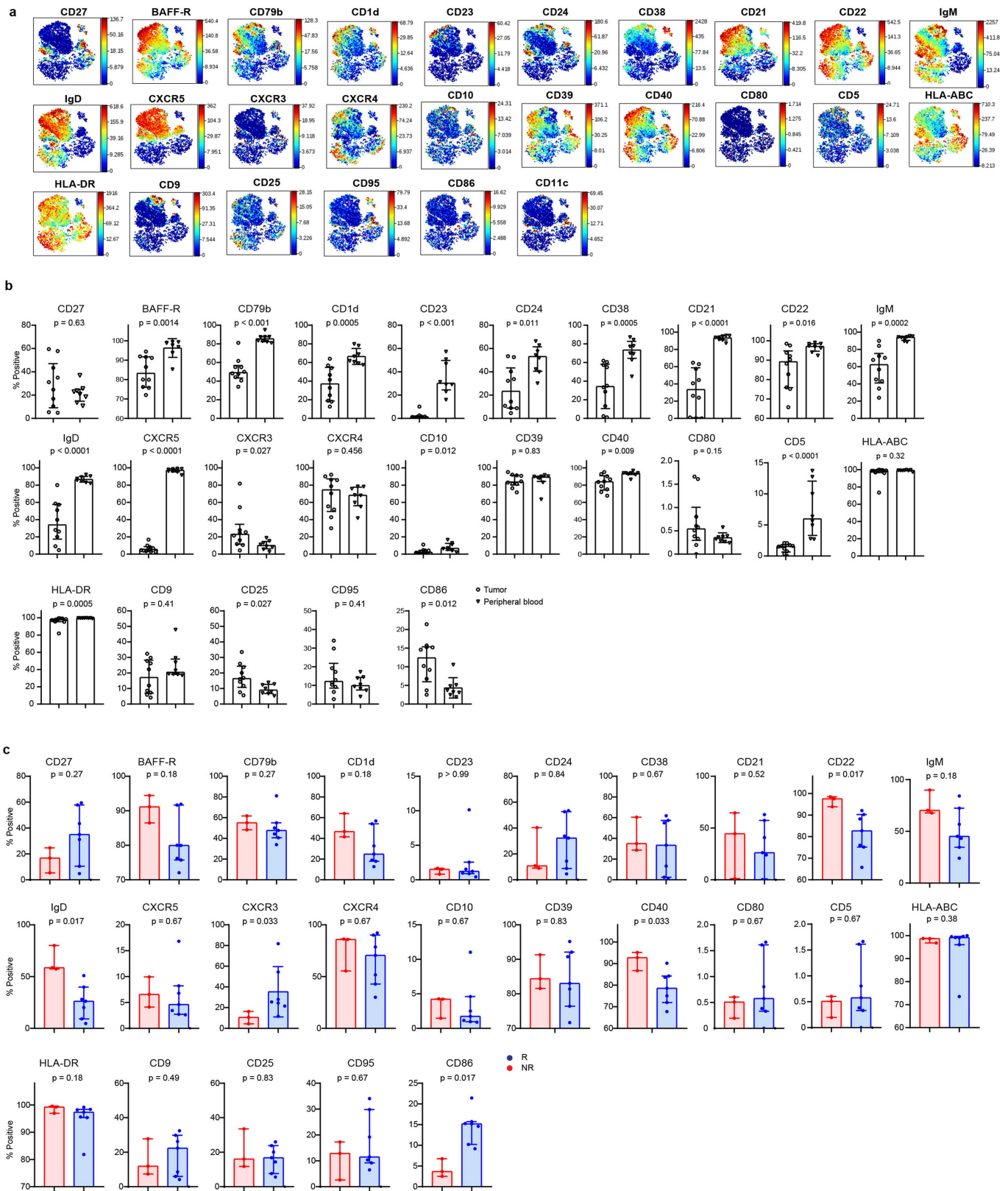
of discriminative genes from all B cells between responder (blue) and non-responder (red) samples. Top marker genes are shown for each group. **c,** Heat map showing scaled expression values ( $\log_2(\text{TPM} + 1)$ ) of discriminative gene sets per cluster as defined in Fig. 3c. A list of representative genes is shown per cluster next to the left margin. For both heat maps, colour scheme is based on z-scores from  $-2.5$  (blue) to  $+2.5$  (yellow).



**Extended Data Fig. 10 | Mass cytometry reveals significant differences in B cell populations between responder and non-responder tumours. a.** Pie charts representing composition of individual tumour and peripheral blood samples for patients with melanoma treated with ICB used in all analyses for mass cytometry. Matched patient samples are located directly beneath one another. Samples from patients with lymph node or non-lymph node metastases as indicated. Cell types as indicated. Asterisk indicates samples included in *t*-SNE plots and pie charts in **c**, Fig. 3d–f, and phenographs in Extended Data Fig. 11. **b.** Scatter plots demonstrating quantification of different peripheral blood and intratumoral B cell phenotypes. Median and interquartile range are shown. All samples are represented in **b** (for tumour,

$n = 7$  R and 3 NR and, for peripheral blood,  $n = 4$  R and 4 NR). *P* values were determined by one-sided Mann–Whitney *U*-test. **c.** *t*-SNE plots demonstrating intratumoral B cell phenotypes from the neoadjuvant ICB trial in patients with advanced melanoma grouped by response and including further breakdown of memory cell subtypes and germinal centre B cells. Plots represent combined analyses of tumours ran simultaneously with the peripheral blood samples ( $n = 5$  R and 3 NR) and include baseline and on-treatment samples as detailed in Supplementary 31. **d.** Quantification of B cell subtypes in tumour from mass cytometric analyses in responders and non-responder from all tumours ( $n = 7$  R and 3 NR). Median and interquartile range are shown. *P* values were determined by one-sided Mann–Whitney *U*-test.





**Extended Data Fig. 11 | Surface expression of markers analysed by mass cytometry. a**, Individual phenographs for surface expression of each marker analysed as indicated. These data represent combined tumour and peripheral blood samples from patients with melanoma treated with ICB ran together (8 tumour  $n = 5$  R and  $n = 3$  NR and 8 peripheral blood samples  $n = 4$  R and  $n = 4$  NR), thus eliminating batch effect. **b**, Percentage of  $CD45^+CD19^+$  cells by tissue type—peripheral blood versus tumour—that are positive for each of the

surface markers indicated. **c**, Percentage of  $CD45^+CD19^+$  cells in tumour by response—responder versus non-responder—that are positive for each of the surface markers indicated. For **b** and **c**, all samples are represented (for tumour,  $n = 7$  R and 3 NR and, for peripheral blood,  $n = 4$  R and  $n = 4$  NR). Error bars indicate median and interquartile range. *P* values were determined by two-sided Mann–Whitney *U*-test.



## Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

### Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a Confirmed

- ☐ ☒ The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- ☐ ☒ An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- ☐ ☒ The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- ☐ ☒ A description of all covariates tested
- ☐ ☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- ☐ ☒ A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- ☐ ☒ For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give  $P$  values as exact values whenever suitable.*
- ☒ ☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- ☐ ☒ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- ☒ ☐ Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated
- ☐ ☒ Clearly defined error bars  
*State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on [statistics for biologists](#) may be useful.

### Software and code

Policy information about [availability of computer code](#)

Data collection

Excel spreadsheet.

Data analysis

FastQC (v0.11.5), STAR 2-pass alignment (v2.5.3), RNA-SeQC (v1.1.8), HTSeq-count (v0.9.1) tool  
Affymetrix Transcriptome Analysis Console (TAC, v4.0) software  
R package MCP-counter algorithm (v. 1.1.0)  
modified TRUST algorithm R package tcR (version 3.4.1)  
Deseq2 (v3.6)  
Graphpad version 7  
R software  
nCounter Analysis System  
R package survival  
Aperio Imagescope  
Flowjo version 10.2, Fluidigm normalization software 2, Cytobank

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

## Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The authors declare that the data supporting the findings of this study are available within the paper and its supplementary information files.

## Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/authors/policies/ReportingSummary-flat.pdf](https://www.nature.com/authors/policies/ReportingSummary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample sizes for each test is described in the manuscript and based on the availability of adequate samples for analysis. Mann Whitney U tests were performed to see if significant differences by response by were noted with available sample size.
Data exclusions	No data were excluded.
Replication	No replication was done on individual patient specimens given limited tissue availability. However, the analyses were conducted in multiple cohorts to provide strength to the findings.
Randomization	Patients were dichotomized into responders and non-responders based on clinical responses as detailed in manuscript. Available samples were chosen for analyses to allow for statistical comparisons by response.
Blinding	Not applicable

## Reporting for specific materials, systems and methods

### Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Unique biological materials
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input type="checkbox"/>	<input checked="" type="checkbox"/> Human research participants

### Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input type="checkbox"/>	<input checked="" type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

## Antibodies

Antibodies used	Included in Materials and Methods in tabular form for CyTOF and within experimental details for IHC and also for analyses of exosomes as well as Digital Spatial profiling
Validation	Describe the validation of each primary antibody for the species and application, noting any validation statements on the manufacturer's website, relevant citations, antibody profiles in online databases, or data provided in the manuscript.

## Human research participants

Policy information about [studies involving human research participants](#)

### Population characteristics

Extended data tables 1, 2, 3, and 8 describes the cohort characteristics for this study.

### Recruitment

Patients for this the studies in this manuscript were recruited from the MD Anderson Melanoma Medical Oncology and Surgical Oncology clinics, MD Anderson Genitourinary Medical Oncology clinics, the Netherlands Cancer Institute oncology clinics, and Dana-Farber/Harvard Cancer Center.

## Flow Cytometry

### Plots

Confirm that:

- ☐ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- ☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- ☒ All plots are contour plots with outliers or pseudocolor plots.
- ☒ A numerical value for number of cells or percentage (with statistics) is provided.

### Methodology

#### Sample preparation

Peripheral blood mononuclear cells (PBMCs) and tumor cells were harvested and washed twice with wash buffer (0.5% bovine serum albumin (BSA) in PBS). To determine the live population, cells were stained with cisplatin 1 $\mu$ M for 3 minutes. The reaction was stopped with FACS buffer (2% Fetal Bovine Serum (FBS) in PBS), and the cells were washed once with wash buffer. Cells were then incubated with 5  $\mu$ l of Fc receptor blocking buffer reagent (Miltenyi) for 10 minutes at room temperature. Cells were incubated with surface antibodies at room temperature for 60 minutes, washed twice with wash buffer and stored overnight in 1ml of 1.6% paraformaldehyde (EMD Biosciences) in PBS with 125 nM iridium nucleic acid intercalator (Fluidigm). The next day, samples were washed twice with cell staining buffer, re-suspended in 1 ml of MilliQ dH<sub>2</sub>O, filtered through a 35  $\mu$ m nylon mesh (cell strainer cap tubes, BD, San Jose, CA) and counted. Before analysis, samples were resuspended in MilliQ dH<sub>2</sub>O supplemented with EQTM four element calibration beads at a concentration of 0.5x10<sup>5</sup>/ml.

#### Instrument

Samples were acquired at 300 events/second on a Helios instrument (Fluidigm) using the Helios 6.5.358 acquisition software (Fluidigm).

#### Software

Mass cytometry data were normalized based on EQTM four element signal shift over time using Fluidigm normalization software 2. Initial data processing was performed using Flowjo version 10.2. Mass cytometry data were normalized based on EQTM four element signal shift over time using Fluidigm normalization software 2. Initially, all R and NR normalized FCS files were either concatenated or separately exported for downstream analyses. Data were processed and analyzed using Cytobank.

#### Cell population abundance

Percentages of different sub-populations of B-cells were measured in aggregated R and NR PBMC and tumor samples; statistical analyses performed via unpaired Student's t-test.

#### Gating strategy

CD19+ sample 'clean-up' was performed by gating on intact (191Ir+ DNA stain), no beads (140Ce-), live (198Pt-), no T-cells CD3- (194Pt), no monocytes CD14- (173Yb) and CD45+ (89Y), no NK Cells CD16- (209Bi), CD19+ B-cells. Mass cytometry complex data were analyzed using viSNE, in combination with heat map, to identify distinct subpopulations using the following parameters: CD19(142Nd), CD20(147Sm), CD5(143Nd), HLA-ABC(144Nd), IgD(146Nd), PDL-1(148Nd), HLA-DR(149Sm), CD25(150Nd), IgM(151Eu), CD95(152Sm), CXCR5(153Eu), CD86(154Sm), CD27(155Gd), CXCR3(156Gd), CD10(158Gd), CD39(160Gd), BAFFR(161Dy), CD79b(162Dy), CD1d(163Dy), CD23(164Dy), CD40(165Ho), CD24(166Er), CD38(167Er), CD9(171Yb), CD11c(172Yb), CXCR4(175Lu), and CD22(176Yb).

- ☐ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

# B cells are associated with survival and immunotherapy response in sarcoma

<https://doi.org/10.1038/s41586-019-1906-8>

Received: 29 June 2018

Accepted: 26 November 2019

Published online: 15 January 2020

Florent Petitprez<sup>1,2,3,4</sup>, Aurélien de Reyniès<sup>4,24</sup>, Emily Z. Keung<sup>5,24</sup>, Tom Wei-Wu Chen<sup>6,7,8,9</sup>, Cheng-Ming Sun<sup>1,2,3</sup>, Julien Calderaro<sup>1,10,11</sup>, Yung-Ming Jeng<sup>9,12</sup>, Li-Ping Hsiao<sup>7</sup>, Laetitia Lacroix<sup>1,2,3</sup>, Antoine Bougouin<sup>1,2,3</sup>, Marco Moreira<sup>1,2,3</sup>, Guillaume Lacroix<sup>1,2,3</sup>, Ivo Natario<sup>1,2,3</sup>, Julien Adam<sup>13</sup>, Carlo Lucchesi<sup>14,15</sup>, Yec'han Laizet<sup>14,15</sup>, Maud Toulmonde<sup>14,16</sup>, Melissa A. Burgess<sup>17</sup>, Vanessa Bolejack<sup>18</sup>, Denise Reinke<sup>19</sup>, Khalid M. Wani<sup>20</sup>, Wei-Lien Wang<sup>20</sup>, Alexander J. Lazar<sup>20,21</sup>, Christina L. Roland<sup>5</sup>, Jennifer A. Wargo<sup>5,21</sup>, Antoine Italiano<sup>14,16,22</sup>, Catherine Sautès-Fridman<sup>1,2,3</sup>, Hussein A. Tawbi<sup>23\*</sup> & Wolf H. Fridman<sup>1,2,3\*</sup>

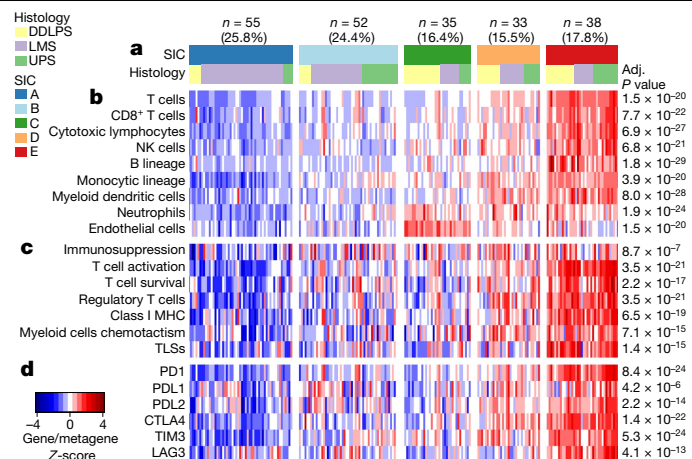
Soft-tissue sarcomas represent a heterogeneous group of cancer, with more than 50 histological subtypes<sup>1,2</sup>. The clinical presentation of patients with different subtypes is often atypical, and responses to therapies such as immune checkpoint blockade vary widely<sup>3,4</sup>. To explain this clinical variability, here we study gene expression profiles in 608 tumours across subtypes of soft-tissue sarcoma. We establish an immune-based classification on the basis of the composition of the tumour microenvironment and identify five distinct phenotypes: immune-low (A and B), immune-high (D and E), and highly vascularized (C) groups. In situ analysis of an independent validation cohort shows that class E was characterized by the presence of tertiary lymphoid structures that contain T cells and follicular dendritic cells and are particularly rich in B cells. B cells are the strongest prognostic factor even in the context of high or low CD8<sup>+</sup> T cells and cytotoxic contents. The class-E group demonstrated improved survival and a high response rate to PD1 blockade with pembrolizumab in a phase 2 clinical trial. Together, this work confirms the immune subtypes in patients with soft-tissue sarcoma, and unravels the potential of B-cell-rich tertiary lymphoid structures to guide clinical decision-making and treatments, which could have broader applications in other diseases.

Soft-tissue sarcomas (STSs) comprise many histological subtypes with distinct clinical and biological behaviours. Genetically ‘simple’ STSs are characterized by translocations that result in fusion proteins and few, if any, other genomic lesions, whereas ‘complex’ STSs have an unbalanced karyotype and several genomic aberrations<sup>1</sup>. STSs are considered ‘non-immunogenic’ with a low mutational burden<sup>2</sup>. Among complex tumours, undifferentiated pleomorphic sarcoma (UPS), dedifferentiated liposarcoma (DDLPS) and—to a lesser extent—leiomyosarcoma (LMS) can exhibit durable responses to immune-checkpoint blockade, whereas simple tumours do not respond to PD1 monotherapy or a combination of anti-PD1 and anti-CTLA4 antibodies<sup>3,4</sup>. Few reports investigating the composition of the tumour microenvironment (TME) in different STS histologies have been published<sup>5–7</sup>, but a recent study

from The Cancer Genome Atlas (TCGA) consortium suggested an association with prognosis<sup>8</sup>.

Here, we developed a new classification of STS, based on the composition of the TME in large cohorts of STS, using the microenvironment cell populations (MCP)-counter method<sup>9</sup>. We found that the B lineage signature—a hallmark of an immune-high class we called E—correlated with an improved survival of patients with STS, in tumours with both high or low infiltration of CD8<sup>+</sup> T cells. In an independent cohort, we used immunohistochemistry to validate the high density of B cells and presence of tertiary lymphoid structures (TLS) in class E. Finally, we showed that class E exhibited the highest response rate to PD1 blockade therapy and improved progression-free survival in a multicentre phase 2 clinical trial of pembrolizumab in STS (SARC028)<sup>4,10</sup>.

<sup>1</sup>Team Cancer, Immune Control and Escape, Centre de Recherche des Cordeliers, INSERM, Paris, France. <sup>2</sup>Centre de Recherche des Cordeliers, Université de Paris, Sorbonne Paris Cité, Paris, France. <sup>3</sup>Centre de Recherche des Cordeliers, Sorbonne University, Paris, France. <sup>4</sup>Programme Cartes d'Identité des Tumeurs, Ligue Nationale Contre le Cancer, Paris, France. <sup>5</sup>Department of Surgical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. <sup>6</sup>Graduate Institute of Oncology, National Taiwan University College of Medicine, Taipei, Taiwan. <sup>7</sup>Department of Oncology, National Taiwan University Hospital, Taipei, Taiwan. <sup>8</sup>National Taiwan University Cancer Center, Taipei, Taiwan. <sup>9</sup>Centers of Genomic and Precision Medicine, National Taiwan University, Taipei, Taiwan. <sup>10</sup>Département de Pathologie, Assistance Publique Hôpitaux de Paris, Groupe Hospitalier Henri Mondor, Creteil, France. <sup>11</sup>Institut Mondor de Recherche Biomédicale, Creteil, France. <sup>12</sup>Department of Pathology, National Taiwan University, Taipei, Taiwan. <sup>13</sup>Department of Biology and Pathology, Gustave Roussy, Villejuif, France. <sup>14</sup>Institut Bergonié, Bordeaux, France. <sup>15</sup>Bioinformatics Unit, Institut Bergonié, Bordeaux, France. <sup>16</sup>Department of Medical Oncology, Institut Bergonié, Bordeaux, France. <sup>17</sup>Department of Medicine, Division of Hematology/Oncology, University of Pittsburgh, Pittsburgh, PA, USA. <sup>18</sup>Cancer Research and Biostatistics, Seattle, WA, USA. <sup>19</sup>Sarcoma Alliance for Research Through Collaboration, Ann Arbor, MI, USA. <sup>20</sup>Department of Pathology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. <sup>21</sup>Department of Genomic Medicine, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. <sup>22</sup>University of Bordeaux, Bordeaux, France. <sup>23</sup>Department of Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. <sup>24</sup>These authors contributed equally: A. de Reyniès, E. Z. Keung. \*e-mail: htawbi@mdanderson.org; herve.fridman@crc.jussieu.fr



**Fig. 1 | The SICs exhibit strongly different TMEs.** This figure refers to the TCGA SARC cohort ( $n = 213$ ). **a**, Composition of the TCGA SARC cohort by SIC, and histology. **b**, Composition of the TME by SIC as defined by the MCP-counter Z-scores. NK cells, natural killer cells. **c**, Expression of gene signatures related to the functional orientation of the immune TME by SIC. **d**, Expression of genes related to immune checkpoints by SIC. Adjusted  $P$  values are obtained from Benjamini-Hochberg correction of two-sided Kruskal-Wallis tests  $P$  values.

## Immune classification of STS

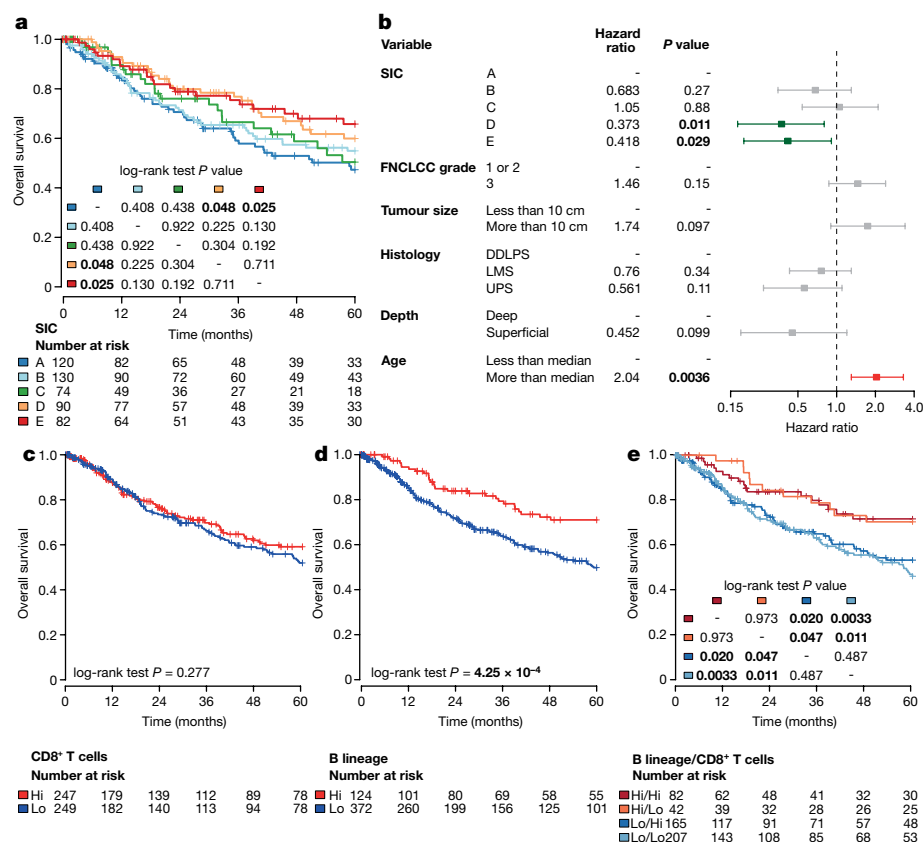
The TME compositions from four independent discovery primary STS datasets (TCGA SARC, Gene Expression Omnibus (GEO) accessions GSE21050, GSE21122 and GSE30929) (Extended Data Table 1) with publicly available gene expression profiles were analysed by MCP-counter, a gene-expression-based TME deconvolution tool<sup>9</sup>. An immune-based classification of STS was developed from this analysis (Extended Data Fig. 1, Methods) and tumours were assigned to one of five sarcoma immune classes (SICs), labelled A, B, C, D and E, with highly distinct profiles (Fig. 1).

We compared the SIC distribution across histological subtypes and found that most LMS tumours were classified to SICs A and B (Fig. 1a). DDLPS accounted for half of SIC C tumours. Tumours classified as SICs D and E were more evenly distributed across histological subtypes. Application of the predictor of the immune classes (Methods) to other STS histologies from French Sarcoma Group (FSG) cohort (Extended Data Table 1) revealed that all SICs could be identified in each histology (Extended Data Fig. 2a).

The TME composition differs significantly between SICs (Fig. 1b). Three SICs showed homogeneous profiles. SIC A, 'immune desert', was characterized by the lowest expression of gene signatures related to immune cells, as well as low vasculature. SIC C, 'vascularized', was dominated by a high expression of endothelial-cell-related genes. SIC E, 'immune and TLS high', was characterized by the highest expression of genes specific to immune populations such as T cells, CD8<sup>+</sup> T cells, natural killer cells and cytotoxic lymphocytes. Notably, a key determinant of SIC E was the high expression of the B lineage signature ( $P = 1.8 \times 10^{-29}$ ). SICs B and D were characterized by heterogeneous but generally immune-low and immune-high profiles, respectively.

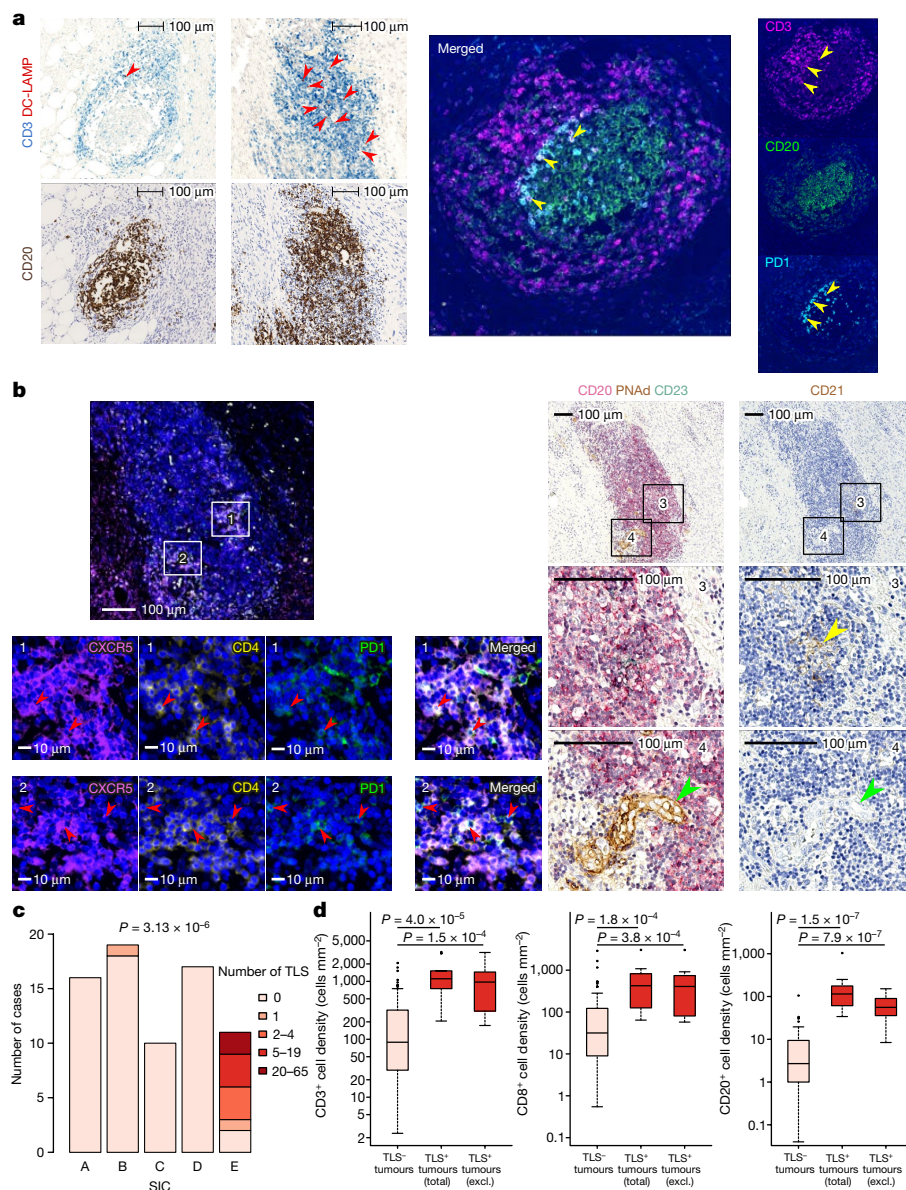
The expression of genes associated with T cell or myeloid cell chemotaxis, T cell activation and survival, major histocompatibility complex class I, and regulatory gene signatures was high in SICs D and E, intermediate in SICs B and C, and very low in SIC A (Fig. 1c). Expression of the lymphoid-structures-associated B-cell-specific chemokine CXCL13 was notably high in E tumours, moderate in D tumours, generally low in B and C tumours, and negligible in A tumours.

The expression of immune-checkpoint-related genes (Fig. 1d) followed that of immune infiltrates, with high expression of the genes encoding PD1, PDL2, CTLA4 and TIM3 (*PDCD1*, *PDCD1LG2*, *CTLA4* and *HAVCR2*, respectively) in SIC E followed by SIC D tumours, and low-to-very-low expression in SIC C, B and A tumours. *CD274* (which encodes PDL1) was heterogeneously expressed across SICs, whereas *LAG3* was expressed at high levels only in SIC E tumours, and its expression was low in all other classes. The above findings were consistent across the four discovery cohorts (Extended Data Fig. 3).



**Fig. 2 | SICs and B cells are predictive of the survival of patients with STS.** This figure refers to TCGA SARC and GSE21050 pooled cohorts ( $n = 496$ ). **a**, Overall survival of patients with STS by the SIC of their tumour. **b**, Multivariate Cox proportional regression outcome, with all included variables represented. For each variable, the reference level is the first one. A grey bar indicates  $P > 0.05$ ; and variables indicated by green and red bars are positively and negatively, respectively, significantly associated with prognosis in this multivariate model. Error bars represent the 95% confidence interval. FNCLCC, Fédération Nationale des Centres de Lutte Contre le Cancer. **c**, **d**, Overall survival of patients with STS according to MCP-counter scores for CD8<sup>+</sup> T cells (**c**) or B lineage cells (**d**). **e**, Overall survival of patients based on the tumour-infiltrating B lineage cells and CD8<sup>+</sup> T cells. Analyses were performed with Kaplan-Meier estimates and two-sided log-rank tests. In each cohort, tumours were considered high (Hi) for CD8<sup>+</sup> T cells if their score was above the median, and high for cytotoxic lymphocytes and B lineage if their score was above the third quartile.





### Fig. 3 | TLSs are a distinguishing feature of the immune-high class of STS.

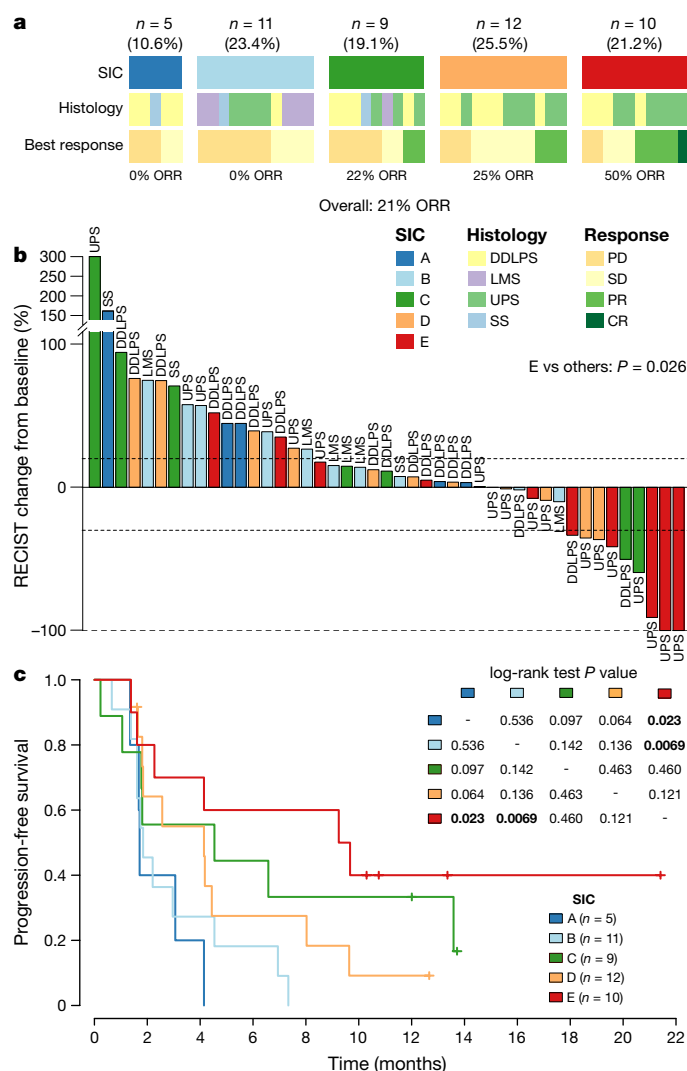
This figure refers to the NTUH cohort ( $n = 93$ ). **a**, Populational characterization of TLSs. Left, examples of two tertiary lymphoid structures by immunohistochemistry, identified as CD3<sup>+</sup> T cell (blue) aggregates containing DC-LAMP<sup>+</sup> mature dendritic cells (red, red arrows) and juxtaposing CD20<sup>+</sup> B cell aggregates (brown). Right, representative immunofluorescence staining of a TLS for CD3 (magenta), CD20 (green) and PD1 (cyan). DAPI staining is shown in blue. The multispectral image shows CD3<sup>+</sup>PD1<sup>+</sup> double-positive cells (yellow arrows). **b**, Functionality of TLSs. Left, CXCR5<sup>+</sup> (magenta), CD4<sup>+</sup> (yellow) and PD1<sup>+</sup> (green) cells in zones 1 and 2 of the same TLS. Multispectral fluorescence images of zones 1 and 2 show CXCR5<sup>+</sup>CD4<sup>+</sup>PD1<sup>+</sup> triple positive cells (red arrows) characteristic of T follicular helper cells. Right, CD20<sup>+</sup> cells stained in pink (left) on consecutive sections of a TLS. CD23 (green on left) and CD21 (brown on

right) positive cells with reticular morphology characteristic of follicular dendritic cells (yellow arrow, zone 3). PNAD<sup>+</sup> structures (brown, green arrow) with high endothelial venule morphology are also detectable nearby (zone 4). **c**, Number of TLS among 5 SICs of 73 tumours of NTUH cohort ( $n = 73$ ). **d**, Characterization of the immune infiltrate in tumours according to TLS presence (TLS<sup>-</sup>  $n = 82$ , TLS<sup>+</sup>  $n = 11$ , total  $n = 93$ ). Densities of CD3<sup>+</sup> (left), CD8<sup>+</sup> (centre) and CD20<sup>+</sup> (right) cells in tumours lacking or containing TLSs; densities including (total) or excluding (excl.) TLS are indicated for the TLS<sup>+</sup> tumours. Box plots represent median (larger bar) and interquartile range (IQR). Upper whisker extends to whichever is minimal, maximum or third quartile plus  $1.5 \times$  IQR. Lower whisker extends to whichever is maximal, minimum or first quartile minus  $1.5 \times$  IQR.  $P$  values were determined by chi-squared test (c) or two-sided Mann-Whitney tests (d).

### SICs are associated with patient survival

After confirmation that the two cohorts with available survival data (TCGA SARC,  $n = 213$ ; GSE21050,  $n = 283$ ) exhibited similar survival patterns (data not shown), the cohorts were pooled to study the clinical outcome of the five SICs (Fig. 2a). Patients with SIC A exhibited the shortest overall survival compared with group D or E patients ( $P = 0.048$  and  $P = 0.025$ , respectively). Similarly, among the other STS histologies from the FSG cohort, patients with SIC A had a shorter overall survival than patients with SIC E (Extended Data

Fig. 2b). In a multivariate model with classical prognostic factors (Fig. 2b), SICs were found to be significantly associated with prognosis, independent of other clinical parameters (as compared with SIC A;  $P = 0.011$  and  $P = 0.029$ , for SICs D and E, respectively). Tumours were separated between high and low expression of CD8<sup>+</sup> T cells, cytotoxic lymphocytes and B lineage signatures based on the observation of the MCP-counter scores distribution (Extended Data Fig. 4). Detailed analysis of the effect of these immune cell population signatures revealed that whereas neither CD8<sup>+</sup> T cells ( $P = 0.277$ ) (Fig. 2c) nor cytotoxic lymphocytes ( $P = 0.0513$ ) (Extended Data



**Fig. 4 | SICs are strongly associated with STS response to PD1 blockade therapy.** This figure refers to the SARC028 cohort (n = 47). **a**, Relationship between SIC, histology and response to treatment in the SARC028 cohort. **b**, Waterfall plot showing the best response to pembrolizumab as a percentage change in the size of target lesions from baseline (n = 45). Tumour sizes were calculated as the sum of target lesion diameters. Colours indicate the SIC to which each tumour was assigned. Dashed lines indicate +20%, -30% and -100% change from baseline levels. SIC E versus other comparison was performed using a two-sided Mann-Whitney test. CR, complete response; PD, progressive disease; PR, partial response; SD, stable disease; SS, synovial sarcoma. **c**, Progression-free survival of patients by tumour SIC (n = 47).

Fig. 5a) significantly correlated with survival, the B lineage signature was significantly associated with improved overall survival ( $P = 4.25 \times 10^{-4}$ ) (Fig. 2d). When analysed in the context of high or low infiltration of CD8<sup>+</sup> T cells (Fig. 2e), cytotoxic lymphocytes or the expression of *PDCD1* (PD1), *CD274* (PDL1) or *FOXP3* (Extended Data Fig. 5b–e), the B lineage signature was the dominant parameter for improved survival, regardless of the expression of other immune factors. In addition, SIC E tumours demonstrate high expression of both *IGJ* (also known as *JCHAIN*) and *TNFRSF17* (encoding BCMA) (data not shown), which indicates that plasma cells<sup>11</sup> may contribute to improved prognosis.

## Mutational landscape of SICs in TCGA SARC

The overall tumour mutational burden was low across the studied cohorts (median: 32 non-synonymous mutations) and appeared to

be similar across all SICs (Extended Data Fig. 6a). However, a few highly mutated tumours (each with more than 250 non-synonymous mutations) were found in the D and E groups. Qualitative mutational analysis revealed several commonly mutated genes across the cohort, including *TP53* (35.2%), *ATRX* (16.0%), *TTN* (9.9%), *RB1* (8.9%), *MUC16* (8.0%), *PCLO* (6.1%), *DNAH5*, *MUC17* and *USH2A* (5.2% each) (Extended Data Fig. 6b). *TP53* was more frequently mutated among SICs D and E tumours ( $P = 0.01$ ) (Extended Data Fig. 6c).

The landscape of copy-number variations, assessed on the TCGA SARC cohort, revealed differences between histologies, consistent with previous observations<sup>8</sup>. However, there was no notable difference in copy-number variation between SICs (data not shown).

## In situ validation of SIC profiles in tumours

To validate the TME profiles of SICs in situ, we analysed an independent cohort of 93 STS cases (NTUH cohort) (Extended Data Table 1). Seventy-three samples passed quality control for transcriptomic analysis using Nanostring nCounter technology. We classified this cohort into the same five SICs (Methods) with the following distribution: A, 16 (21.9%); B, 19 (26.0%); C, 10 (13.7%); D, 17 (23.3%); and E, 11 (15.1%). The NTUH cohort samples exhibited gene-expression-based TME profiles that were similar to that of TCGA SARC and GSE21050 cohorts (Extended Data Fig. 7a).

By quantitative immunohistochemistry, immune-desert SIC A was characterized by very low densities of CD3<sup>+</sup>, CD8<sup>+</sup> or CD20<sup>+</sup> cells, whereas immune-and-TLS-high SIC E exhibited high densities of these cells (pairwise comparison,  $P = 4.01 \times 10^{-6}$ ,  $P = 6.64 \times 10^{-6}$  and  $P = 9.90 \times 10^{-7}$ , respectively). The vascularized SIC C exhibited a moderate infiltration by immune cells and a high density of CD34<sup>+</sup> endothelial cells (Extended Data Fig. 7b, c).

## TLSS are a feature of SIC E tumours

The CXCL13 chemokine, which is associated with the presence of TLSS<sup>12</sup>, was strongly expressed in SIC E tumours (Fig. 1c, Extended Data Fig. 2c). Expression of CXCL13 was highly correlated with that of the TLS-associated 12-chemokine signature<sup>13</sup> (Extended Data Fig. 8a), which suggests that TLSS could be a marker of SIC E. TLSS were defined as a CD20<sup>+</sup> B-cell follicle juxtaposed to a CD3<sup>+</sup> T cell aggregate containing at least one DC-LAMP<sup>+</sup> (also known as LAMP3<sup>+</sup>) mature dendritic cell<sup>12,14–16</sup> (Fig. 3a, left). A strong association between SICs and the presence of TLSS was identified ( $P = 3.13 \times 10^{-6}$ ) (Fig. 3c). No TLSS were observed in tumours from SICs A, C and D, and only one tumour from SIC B had one TLS. By contrast, nine out of eleven (82%) SIC E tumours exhibited one or more TLS. All TLSS were intratumoural (Extended Data Fig. 8b), and found at the periphery and in the centre of the tumour in all histologies (Extended Data Fig. 8c, d).

We observed the presence of CD3<sup>+</sup>PD1<sup>+</sup> T cells (Fig. 3a, right) in the germinal centre of TLSS with characteristics of follicular T helper cells<sup>17,18</sup> (positive for CD4, PD1 and the CXCL13 receptor CXCR5) (Fig. 3b, left), CD23<sup>+</sup>CD21<sup>+</sup> cells with reticular morphology characteristic of follicular dendritic cells, and peripheral node addressin (PNA<sup>+</sup>) positive structures with high endothelial venule morphology (Fig. 3b, right). Germinal centres are a hallmark of secondary follicle-like TLSS (SFL-TLS), the final maturation step of TLS; the earlier steps being early TLSS (E-TLS) and primary follicle-like TLSS (PFL-TLS)<sup>15,16</sup>. E-TLS, PFL-TLS and SFL-TLS represented 60.5%, 21.1% and 18.3%, respectively, of all TLSS analysed (Extended Data Fig. 8e, f). This differed between histologies ( $P = 7.76 \times 10^{-5}$ ), with UPS having only 16.7% of E-TLS.

Tumours with TLSS (11.8%, 11 out of 93) had significantly higher densities of tumour-infiltrating CD3<sup>+</sup> T cells ( $P = 4.0 \times 10^{-5}$ ), CD8<sup>+</sup> T cells ( $P = 1.8 \times 10^{-4}$ ) and CD20<sup>+</sup> B cells ( $P = 1.5 \times 10^{-5}$ ) (Fig. 3d). This association persisted even if T and B cells within TLSS were excluded from the analysis ( $P = 1.5 \times 10^{-4}$ ,  $P = 3.8 \times 10^{-4}$  and  $P = 7.9 \times 10^{-7}$ , respectively) (Fig. 3d), which suggests that high immune cell infiltration is not limited to TLSS.

## SICs predict patient response to PD1 blockade

We examined whether SICs can predict the patient response to checkpoint blockade therapy. We obtained 47 pre-treatment STS metastasis biopsies from patients enrolled in the SARCO28 clinical trial<sup>4</sup> and its expansion cohort<sup>10</sup> (Extended Data Table 1), which evaluated the efficacy of the anti-PD1 monoclonal antibody pembrolizumab in patients with metastatic STS. Of these 47 patients, 1 achieved a complete response, 9 a partial response, 17 stable disease and 20 had progressive disease (Fig. 4a). Pre-treatment tumours were classified into SICs based on gene-expression data. The objective response rate (ORR) (which accounts for complete and partial responses) as evaluated by response evaluation criteria in solid tumours (RECIST) criteria was 21.2% in the overall cohort. SICs, however, showed substantial variation in ORR, with SIC E patients exhibiting the highest ORR (50%, 5 out of 10), followed by SIC D (25%, 3 out of 12) and SIC C (22%, 2 out of 9) (Fig. 4a). A complete response was found only in SIC E, as well as one patient who had a 100% change in target lesions but a non-complete response in non-target lesions and thus did not qualify for a complete response. Notably, there were no responders within the SIC A (0 out of 5) and B (0 out of 11) groups (Fig. 4a). Overall, SIC E tumours were associated with the highest response rate to pembrolizumab in comparison with tumours from other SICs ( $P = 0.026$ , Fig. 4b). Patients with SIC E tumours also exhibited improved progression-free survival compared with patients with SIC A or B tumours ( $P = 0.023$  and  $P = 0.0069$ , respectively) (Fig. 4c).

## Discussion

This study is, to our knowledge, the most comprehensive analysis of the STS immune TME and the first to evaluate the prognostic effect of immune infiltrates by simultaneously integrating several immune cell populations and malignant cell characteristics. Previous studies have examined the immune profile of STS tumours, but the importance of B cells and TLSSs was not investigated. The clinical effect of CD8<sup>+</sup> T cells and PD1 expression has yielded controversial results<sup>7,8,19–25</sup>. Here, we found the CD8<sup>+</sup> T cell signature and PD1 were expressed in class D and E SICs, which are associated with favourable outcomes, providing high infiltration of B cells. The integrative analysis demonstrates that infiltration by B cells is a key discriminative feature of a group of patients with improved survival. This B-cell-high group was found to respond better to PD1 blockade therapy, although this should be validated on a larger cohort.

The field of immuno-oncology is rapidly expanding, and is crucial to accurately identify patients who are likely to respond. Here, we propose a classification for STS that is immune-centric with prognostic effect. It defines a group of patients with a better response to anti-PD1 therapy marked by B cells and TLSSs. This finding may have broad applications. Sarcomas are considered immune-quiescent tumours, with a low mutational burden. Nevertheless, our data show that some STSs are immunogenic and that this is driven by B cells. Further work is needed to extend these findings to all STS histologies and other cancers. Similarly, the underlying mechanisms require further investigation, but a possible explanation is that TLSSs are sites at which anti-tumoral immunity is generated, with B cells instructing T cells—in particular CD8<sup>+</sup> T cells—to recognize tumour-associated antigens<sup>26</sup>. It is noteworthy that TLS-rich tumours are more infiltrated by CD8<sup>+</sup> T cells. These T cells can become exhausted, explaining the correlation of the expression of immune checkpoints (such as PD1 and LAG3) with TLSSs, and why treatment with checkpoint inhibitors may allow productive anti-tumour immunity in TLS-rich tumours. Overall, our findings lay the foundation for a tool to risk-stratify patients with STS and identify those who may be more likely to benefit from immunotherapies, and may be broadly applicable to other malignancies<sup>26–30</sup>.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-019-1906-8>.

1. Helman, L. J. & Meltzer, P. Mechanisms of sarcoma development. *Nat. Rev. Cancer* **3**, 685–694 (2003).
2. Fletcher, C., Bridge, J., Hogendoorn, P. & Mertens, F. *WHO Classification of Tumours of Soft Tissue and Bone* (World Health Organization, 2013).
3. D'Angelo, S. P. et al. Nivolumab with or without ipilimumab treatment for metastatic sarcoma (Alliance A091401): two open-label, non-comparative, randomised, phase 2 trials. *Lancet Oncol.* **19**, 416–426 (2018).
4. Tawbi, H. A. et al. Pembrolizumab in advanced soft-tissue sarcoma and bone sarcoma (SARCO28): a multicentre, two-cohort, single-arm, open-label, phase 2 trial. *Lancet Oncol.* **18**, 1493–1501 (2017).
5. Beck, A. H. et al. Discovery of molecular subtypes in leiomyosarcoma through integrative molecular profiling. *Oncogene* **29**, 845–854 (2010).
6. Gibault, L. et al. New insights in sarcoma oncogenesis: a comprehensive analysis of a large series of 160 soft tissue sarcomas with complex genomics. *J. Pathol.* **223**, 64–71 (2011).
7. Pollack, S. M. et al. T-cell infiltration and clonality correlate with programmed cell death protein 1 and programmed death-ligand 1 expression in patients with soft tissue sarcomas. *Cancer* **123**, 3291–3304 (2017).
8. Cancer Genome Atlas Research Network. Comprehensive and integrated genomic characterization of adult soft tissue sarcomas. *Cell* **171**, 950–965.e28 (2017).
9. Becht, E. et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol.* **17**, 218 (2016).
10. Burgess, M. A. et al. Clinical activity of pembrolizumab (P) in undifferentiated pleomorphic sarcoma (UPS) and dedifferentiated/pleomorphic liposarcoma (LPS): final results of SARCO28 expansion cohorts. *JCO* **37**, 11015–11015 (2019).
11. Kroeger, D., Milne, K. & Nelson, B. Tumor-infiltrating plasma cells are associated with tertiary lymphoid structures, cytolytic T-cell responses, and superior prognosis in ovarian cancer. *Clinical Cancer Res.* **22**, 3005–3015 (2016).
12. Sautès-Fridman, C., Petitprez, F., Calderaro, J. & Fridman, W. H. Tertiary lymphoid structures in the era of cancer immunotherapy. *Nat. Rev. Cancer* **19**, 307–325 (2019).
13. Coppola, D. et al. Unique ectopic lymph node-like structures present in human primary colorectal carcinoma are identified by immune gene array profiling. *Am. J. Pathol.* **179**, 37–45 (2011).
14. Dieu-Nosjean, M.-C., Goc, J., Giraldo, N. A., Sautès-Fridman, C. & Fridman, W. H. Tertiary lymphoid structures in cancer and beyond. *Trends Immunol.* **35**, 571–580 (2014).
15. Posch, F. et al. Maturation of tertiary lymphoid structures and recurrence of stage II and III colorectal cancer. *Oncol Immunology* **7**, e1378844 (2017).
16. Siliņa, K. et al. Germinal centers determine the prognostic relevance of tertiary lymphoid structures and are impaired by corticosteroids in lung squamous cell carcinoma. *Cancer Res.* **78**, 1308–1320 (2018).
17. Gu-Trantien, C. et al. CD4<sup>+</sup> follicular helper T cell infiltration predicts breast cancer survival. *J. Clin. Invest.* **123**, 2873–2892 (2013).
18. Dorfman, D. M., Brown, J. A., Shahsafaei, A. & Freeman, G. J. Programmed death-1 (PD-1) is a marker of germinal center-associated T cells and angioimmunoblastic T-cell lymphoma. *Am. J. Surg. Pathol.* **30**, 802–810 (2006).
19. D'Angelo, S. P. et al. Prevalence of tumor-infiltrating lymphocytes and PD-L1 expression in the soft tissue sarcoma microenvironment. *Hum. Pathol.* **46**, 357–365 (2015).
20. Sorbye, S. W. et al. Prognostic impact of peritumoral lymphocyte infiltration in soft tissue sarcomas. *BMC Clin. Pathol.* **12**, 5 (2012).
21. Sorbye, S. W. et al. High expression of CD20<sup>+</sup> lymphocytes in soft tissue sarcomas is a positive prognostic indicator. *Oncol Immunology* **1**, 75–77 (2012).
22. Bertucci, F. et al. PDL1 expression is a poor-prognosis factor in soft-tissue sarcomas. *Oncol Immunology* **6**, e1278100 (2017).
23. Kim, J. R. et al. Tumor infiltrating PD1-positive lymphocytes and the expression of PD-L1 predict poor prognosis of soft tissue sarcomas. *PLoS One* **8**, e82870 (2013).
24. Honda, Y. et al. Infiltration of PD-1-positive cells in combination with tumor site PD-L1 expression is a positive prognostic factor in cutaneous angiosarcoma. *Oncol Immunology* **6**, e1253657 (2016).
25. Paydas, S., Bagir, E. K., Deveci, M. A. & Gonlusen, G. Clinical and prognostic significance of PD-1 and PD-L1 expression in sarcomas. *Med. Oncol.* **33**, 93 (2016).
26. Nielsen, J. S. et al. CD20<sup>+</sup> tumor-infiltrating lymphocytes have an atypical CD27<sup>+</sup> memory phenotype and together with CD8<sup>+</sup> T cells promote favorable prognosis in ovarian cancer. *Clin. Cancer Res.* **18**, 3281–3292 (2012).
27. Montfort, A. et al. A strong B-cell response is part of the immune landscape in human high-grade serous ovarian metastases. *Clin. Cancer Res.* **23**, 250–262 (2017).
28. Hennequin, A. et al. Tumor infiltration by Tbet<sup>+</sup> effector T cells and CD20<sup>+</sup> B cells is associated with survival in gastric cancer patients. *Oncol Immunology* **5**, e1054598 (2015).
29. Wouters, M. C. A. & Nelson, B. H. Prognostic significance of tumor-infiltrating B cells and plasma cells in human cancer. *Clin. Cancer Res.* **24**, 6125–6135 (2018).
30. Helmink, B. et al. B cells and tertiary lymphoid structures promote immunotherapy response. *Nature* <https://doi.org/10.1038/s41586-019-1922-8> (2020).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2020



## Methods

### Ethics and patients

Patients diagnosed with DDLPS, LMS and UPS were identified and the pathology diagnosis was confirmed by a certified pathologist in National Taiwan University Hospital. The research was approved by the Research Ethics Committee of NTUH (201605061RINA) for this retrospective study. Formalin-fixed paraffin-embedded (FFPE) blocks were retrieved and 4–5- $\mu$ m-thick slides were taken for immunohistochemistry staining and RNA extraction for Nanostring testing. Other cohorts were previously published<sup>4,8,31–35</sup>.

### Establishing the immune classification of STS

To establish a robust immune classification of STS, publicly available transcriptomic data from TCGA data portal and the GEO repository representing four large and independent patient cohorts were included. Only tumours from the most common histologies of genomically complex STS were included: LMS, UPS and DDLPS. We analysed data from the TCGA SARC<sup>8</sup> ( $n = 213$ ), GSE21050<sup>31</sup> ( $n = 283$ ), GSE21122<sup>32</sup> ( $n = 72$ ) and GSE30929<sup>33</sup> ( $n = 40$ ) cohorts.

### Public transcriptomic data pre-processing

Transcriptomic data were downloaded from the TCGA data portal (SARC cohort) and GEO (accessions GSE21050, GSE21122 and GSE30929). TCGA SARC was restricted to complex genomics sarcomas (UPS, DDLPS and LMS). Normalized TCGA SARC RNA-sequencing data were log<sub>2</sub>-transformed. Microarray data were normalized using frozen-RMA method<sup>36</sup> from the R package *frma*. Batch effect was corrected across series using ComBat<sup>37</sup>, with histology as covariate.

### Estimation of the TME composition

The TME composition of each tumour was assessed with the MCP-counter tool<sup>9</sup>, which provides abundance scores for eight immune (T cells, CD8<sup>+</sup> T cells, cytotoxic lymphocytes, natural killer cells, B cell lineage, monocytic lineage, myeloid dendritic cells and neutrophils), and two stromal populations (endothelial cells and fibroblasts). The scores are based on analysis of transcriptomic markers—that is, transcriptomic features that are strongly, specifically and stably expressed in a unique cell population. These scores are proportional to the abundance of each cell population in the tumour, therefore allowing inter-sample comparison and large cohort analyses<sup>38</sup>. The MCP-counter signatures composition are as follows: T cells: *CD28*, *CD3D*, *CD3G*, *CD5*, *CD6*, *CHRM3-AS2*, *CTLA4*, *FLT3LG*, *ICOS*, *MAL*, *PBX4*, *SIRPG*, *THEMIS*, *TNFRSF25* and *TRAT1*; CD8<sup>+</sup> T cells: *CD8B*, cytotoxic lymphocytes: *CD8A*, *EOMES*, *FGFBP2*, *GNLY*, *KLRC3*, *KLRC4* and *KLRD1*; B lineage: *BANK1*, *CD19*, *CD22*, *CD79A*, *CR2*, *FCRL2*, *IGKC*, *MS4A1* and *PAX5*; natural killer cells: *CD160*, *KIR2DL1*, *KIR2DL3*, *KIR2DL4*, *KIR3DL1*, *KIR3DS1*, *NCR1*, *PTGDR* and *SH2D1B*; monocytic lineage: *ADAP2*, *CSF1R*, *FPR3*, *KYNU*, *PLA2G7*, *RASSF4* and *TFEC*; myeloid dendritic cells: *CD1A*, *CD1B*, *CD1E*, *CLEC10A*, *CLIC2* and *WFDC21P*; neutrophils: *CA4*, *CEACAM3*, *CXCR1*, *CXCR2*, *CYP4F3*, *FCGR3B*, *HAL*, *KCNJ15*, *MEGF9*, *SLC25A37*, *STEAP4*, *TECPR2*, *TLE3*, *TNFRSF10C* and *VNN3*; endothelial cells: *ACVRL1*, *APLN*, *BCL6B*, *BMP6*, *BMX*, *CDH5*, *CLEC14A*, *CXorf36* (also known as *DIPK2B*), *EDN1*, *ELTD1*, *EMCN*, *ESAM*, *ESM1*, *FAM124B*, *HECW2*, *HHIP*, *KDR*, *MMRN1*, *MMRN2*, *MYCT1*, *PALMD*, *PEAR1*, *PGF*, *PLXNA2*, *PTPRB*, *ROBO4*, *SDPR*, *SHANK3*, *SHE*, *TEK*, *TIE1*, *VEPH1* and *VWF*.

### Intracohort immune classifications

The fibroblasts signature was removed from this analysis as all STS tumours exhibited high and homogeneous scores for this cell population, which is consistent with the mesenchymal origin of STS. The signature for CD8 T cells was removed from the analysis for GSE21050, GSE21122 and GSE30929 as it showed very small variation across all samples in these microarray-based cohorts. Unsupervised clustering of samples in each cohort was performed based on the metagene

Z-score for the included populations of MCP-counter (Extended Data Fig. 9a–d) using R software, with the Euclidian distance and Ward's linkage criterion, using the *gplots* package. The TCGA SARC, GSE21050, GSE21122 and GSE30929 cohorts were separated into 6, 9, 7 and 6 groups, respectively. The number of clusters was chosen empirically following the dendrograms shown in Extended Data Fig. 9a–d. Analysis of the intersample variance revealed that much of the explainable variance was already attained at the chosen number of clusters as visualized in Extended Data Fig. 9e–h.

### Pan-cohort immune classes

To aggregate the above four intracohort classifications, the transcriptome matrix of each cohort was independently zero-centred for each gene across all samples. Then, we computed the centroids of each class over the whole transcriptome and analysed the Pearson correlations between all the centroids on the set of genes shared across the four cohorts (Extended Data Fig. 9i). From these correlations, we deduced five SICs. The tumours from six remaining cohort-specific clusters shared intermediate/weak correlation patterns to other clusters and were temporarily labelled as 'unclassified'.

### Prediction of the immune classes

Centroids of SICs were computed on MCP-counter intraseries Z-scores for T cells, cytotoxic lymphocytes, B cell lineage, natural killer cells, monocytic lineage, myeloid dendritic cells, neutrophils and endothelial cells, on all cohorts. To predict de novo the immune classes of each of the cohorts, MCP-counter Z-scores were computed, and each sample was assigned to the closest immune class based on its Euclidian distance to the related centroids. The SICs labels used are the ones predicted using this method. Principal component analysis of the 608 samples on the MCP-counter scores shows that the intra-SIC homogeneity was improved by this prediction step (Extended Data Fig. 9j, k), as confirmed by supervised tests across SICs (Extended Data Fig. 9l, m).

### Gene signatures for the functional orientation

The signatures used to determine the functional orientation of the TME were derived from the literature<sup>39</sup>. The signatures were the following: immunosuppression (*CXCL12*, *TGFB1*, *TGFB3* and *LGALS1*), T cell activation (*CXCL9*, *CXCL10*, *CXCL16*, *IFNG* and *IL15*), T cell survival (*CD70* and *CD27*), regulatory T cells (*FOXP3* and *TNFRSF18*), major histocompatibility complex class I (*HLA-A*, *HLA-B*, *HLA-C*, *HLA-E*, *HLA-F*, *HLA-G* and *B2M*), myeloid cell chemotaxis (*CCL2*), and tertiary lymphoid structures (*CXCL13*). For each signature, scores were computed as the geometric mean signature expression.

### De novo prediction of the immune classes of additional cohorts and other platforms

The predictor described above was adapted to analyse new and independent samples, from Nanostring-analysed FFPE samples. In a first step, SICs were estimated on the NTUH cohort by sorting samples on the B lineage signature, T cells signature then endothelial cell signature and assigning each sample according to the SIC it resembled the most. Similar to as described above, centroids of each SIC on Nanostring data MCP-counter scores Z-scores were computed and samples were reassigned to the SIC they were closest to the centroid of. For new samples from the SARC028 cohort, MCP-counter scores for T cells, cytotoxic lymphocytes, B lineage and endothelial cells were computed and transformed as Z-scores. Distances with Nanostring-defined centroids presented above were computed with Euclidian metric, and samples were assigned to the SIC with the lowest distance.

### RNA extraction from FFPE tumours

Human FFPE tumour specimens were cut into 3- $\mu$ m-thick sections and were reviewed under microscope for tumour histology. Non-tumour tissues were excluded and tumour tissues were deparaffinized by

# Article

deparaffinization solution (Qiagen 19093) and RNA were extracted by RNeasy FFPE kit (Qiagen 73504) according to the manufacturer's protocol. RNA quality and size distribution were determined by the Agilent 2100 Bioanalyzer with RNA analysis kits (RNA 6000 nano kit 5067-1511, RNA 6000 nano reagent 5067-1512, RNA 6000 nano ladder 5067-1529, RNA 6000 pico kit 5067-1513, RNA 6000 pico reagents 5067-1514, RNA 6000 pico ladder 5067-1535) for cohorts NTUH core and NTUH whole, and by the Agilent RNA ScreenTape assay (catalogue: RNA ScreenTape 5067-5576, RNA ScreenTape sample buffer 5067-5577, RNA ScreenTape ladder 5067-5578) and Agilent 2200 TapeStation for cohort SARCO28. The samples from SARCO28 were separately quality-controlled by the sarcoma pathology group at MD Anderson Cancer Center.

## Nanostring nCounter analysis

The RNA was analysed using the nCounter Technology (Nanostring Technologies) as per the manufacturer's protocol. Data were normalized using the nSolver software (Nanostring Technologies).

## Enzymatic and fluorescent multiplexed immunohistochemistry

The FFPE human tumour and control specimens were cut into 3- $\mu$ m-thick sections. Human FFPE tonsil sections were used as positive controls for CD3, CD4, CD8, CD20, CD21, CD23, CD34, CXCR5, DC-LAMP, PD1, PDL1 and PNAd, placenta sections were used in addition for PDL1 and cerebral cortex tissue was used as a negative control. The specificity of all antibodies was tested by the manufacturers and the specificity of anti-PD1 antibodies was validated in our laboratory on overexpressing cells pellets as previously reported<sup>40</sup>. Antigen retrieval was carried out on a PT-link (Dako) using the EnVision FLEX Target Retrieval Solutions at High pH (Dako, K8004) or Low pH (Dako, K8005). Endogenous peroxidase activity and non-specific Fc receptor binding were blocked with H2O2 3% (Giffrer, 10603051) and Protein Block (Dako, X0909) respectively. The primary and secondary antibodies used for immunohistochemistry and immunofluorescence are summarized in Extended Data Table 2. Immunohistochemistry and immunofluorescence images were independently analysed blindly by three observers (L.L., C.S.-F. and G.L.).

## Enzymatic immunohistochemistry

The stainings were performed with an Autostainer Link 48 (Dako). Chromogenic detection was performed using 3,3'-diaminobenzidine (Dako, K3468) for CD8, CD20, CD21, PDL1 and PNAd; 3-amino-9-ethylcarbazole substrate (Vector Laboratories, SK-4200) for DC-LAMP; Blue Alkaline Phosphatase Substrate (Vector Laboratories, SK5300) for CD3; High-Def red IHC chromogen (AP) (Enzo, ADI-950-140-0030) for CD20; and Permanent HRP Green (Zytomed Systems, ZUC070-100) for CD23 and CD34. The nuclei were counterstained with haematoxylin (Dako, S3301). After mounting with Glycergel Mounting Medium (Dako, C056330-2) or EcoMount (Biocare Medical, EM897L), the slides were scanned with a Nanozoomer (Hamamatsu). For CD3, CD8, CD20 and DC-LAMP markers, the density of positive cells per mm<sup>2</sup> was quantified with Calopix Software (Tribvn). For CD34 marker, the density of positive vessels per mm<sup>2</sup> was quantified with Halo10 software (Indica labs). TLS were identified using the registration module to fit one slide on the other (Halo10 software, Indica labs). Tumours were considered TLS-positive when a CD3 aggregate with DC-LAMP staining was found juxtaposing a CD20 aggregate. Only aggregates with surface above 60,000  $\mu$ m<sup>2</sup>, containing at least 700 cells and at least 350 CD20<sup>+</sup> cells were considered.

## Fluorescent multiplexed immunohistochemistry

For the PD1, CD20 and CD3 3-plex staining, a tyramide system amplification (TSA) was used. The stainings were performed with a Leica Bond RX. The incubation with TSA reagent was performed after the incubation of the horseradish peroxidase (HRP)-conjugated polymer and was followed by antibody stripping at 97 °C for 10 min. This protocol was repeated for the second and third primary antibodies and

corresponding polymer incubations. The dilutions used for the TSA are 1:400 for TSA AF488, 1:800 for TSA AF594 and 1:200 for TSA AF647, as per the manufacturer's recommendations. For the CXCR5, CD4 and PD1 3-plex staining, we used a conventional fluorescent-dye conjugated secondary antibody system performed manually (all secondary antibodies were diluted at 1:100). For all the fluorescent stainings, the nuclei were stained with DAPI Solution (Thermo Fisher, 62248) at 2  $\mu$ g ml<sup>-1</sup> for 10 min. After mounting with ProLongTM Gold Antifade Mountant (ThermoFisher, P36934), the slides were scanned with a Zeiss Axio Scan.Z1.

## Statistical analysis

All statistical analyses were performed using the R software (v.3.4.4) and the packages survival, gplots, dunn.test and FactoMineR. The relationship between two categorical variables was estimated with the chi-squared test. The relationship between a categorical variable and a quantitative variable was estimated with the Mann-Whitney *U* test (two categories) or the Kruskal-Wallis test (three or more categories). All tests were two-sided. In cases with three or more categories, pairwise comparisons were carried out with Dunn tests. The relationship between two quantitative variables was estimated with the Pearson correlation. When appropriate, *P* values were corrected for multiple hypothesis testing with the Bonferroni or Benjamini-Hochberg methods, as specified in the text or figure legends. Survival was analysed with Kaplan-Meier estimates and log-rank tests. No statistical methods were used to predetermine sample size. The experiments were not randomized, and investigators were not blinded to allocation during experiments and outcome assessment unless stated otherwise.

## Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

## Data availability

The transcriptomic datasets analysed in this study can be accessed on the GDC Portal ([portal.gdc.cancer.gov](https://portal.gdc.cancer.gov), cohort TCGA SARC) and the GEO repository under accession numbers GSE21050, GSE21122 and GSE30929. FSG cohort data are publicly available from ArrayExpress for gastrointestinal stromal tumour with accession code E-MTAB-373, and from the GEO for synovial sarcomas with accession number GSE40021. Myxoid liposarcomas from the FSG cohort are available from the corresponding authors upon reasonable request. Immunohistochemistry and gene expression data related to the NTUH cohort (Fig. 3, Extended Data Figs. 7, 8) are available upon reasonable request to W.H.F. ([herve.fridman@crc.jussieu.fr](mailto:herve.fridman@crc.jussieu.fr)). The data that support the findings related to Fig. 4 are available from SARC but restrictions apply to the availability of these data, which were used under license for the study. Data are, however, available from H.A.T. ([htawbi@mdanderson.org](mailto:htawbi@mdanderson.org)) upon reasonable request and with permission of SARC.

## Code availability

All code used in this study is available from the corresponding author upon reasonable request.

31. Chibon, F. et al. Validated prediction of clinical outcome in sarcomas and multiple types of cancer on the basis of a gene expression signature related to genome complexity. *Nat. Med.* **16**, 781-787 (2010).
32. Barretina, J. et al. Subtype-specific genomic alterations define new targets for soft-tissue sarcoma therapy. *Nat. Genet.* **42**, 715-721 (2010).
33. Gobbe, R. M. et al. Expression profiling of liposarcoma yields a multigene predictor of patient outcome and identifies genes that contribute to liposarcomagenesis. *Cancer Res.* **71**, 2697-2705 (2011).
34. Lagarde, P. et al. Mitotic checkpoints and chromosome instability are strong predictors of clinical outcome in gastrointestinal stromal tumors. *Clin. Cancer Res.* **18**, 826-838 (2012).
35. Lagarde, P. et al. Chromosome instability accounts for reverse metastatic outcomes of pediatric and adult synovial sarcomas. *J. Clin. Oncol.* **31**, 608-615 (2013).
36. McCall, M. N., Bolstad, B. M. & Irizarry, R. A. Frozen robust multiarray analysis (fRMA). *Biostatistics* **11**, 242-253 (2010).



37. Johnson, W. E., Li, C. & Rabinovic, A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* **8**, 118–127 (2007).
38. Petitprez, F. et al. Transcriptomic analysis of the tumor microenvironment to guide prognosis and immunotherapies. *Cancer Immunol. Immunother.* **67**, 981–988 (2017).
39. Beuselinck, B. et al. Molecular subtypes of clear cell renal cell carcinoma are associated with sunitinib response in the metastatic setting. *Clin. Cancer Res.* **21**, 1329–1339 (2015).
40. Giraldo, N. A. et al. Orchestration and prognostic significance of immune checkpoints in the microenvironment of primary and metastatic renal cell cancer. *Clin. Cancer Res.* **21**, 3031–3040 (2015).

**Acknowledgements** This work was supported by the Institut National de la Santé et de la Recherche Médicale, the Université de Paris, Sorbonne University, the Programme Cartes d'Identité des Tumeurs (CIT) from the Ligue Nationale Contre le Cancer, grants from Institut National du Cancer (HTE-INSERM plan cancer, C16082DS), Association pour la Recherche sur le Cancer (ARC), Cancer Research for Personalized Medicine programme (CARPEM T8), French Sarcoma Group, the European Connective Tissue Cancer Network (CONTICANET, FP6-018806), 'FONCER contre le cancer' programme, Labex Immuno-Oncology (LAXE62\_9UMRS972 FRIDMAN), the National Institutes of Health (E.Z.K. is supported by T32CA0095999) and the Moon Shot program at MD Anderson Cancer Center. Grants from the Ministry of Education (NTU-107L9014) and Ministry of Science and Technology (MOST 107-3017-F-002 -002-), Taiwan and from the National Taiwan University (YongLin Chair Grant S-01 and S-03) also supported this study. SARCO28 was jointly funded by Merck, Inc., SARC, Sarcoma Foundation of America, and the QuadW Foundation. F.P. supported by CARPEM doctorate fellowship. C.L.R. is recipient of the Paul Calabresi Clinical Oncology Award (K12 CA088084). The slides stained for immunofluorescence were scanned and analysed at the Centre d'Histologie, d'Imagerie et de Cytométrie (CHIC), Centre de Recherche des Cordeliers UMR51138 (Paris, France). CHIC is a member of the Sorbonne University Flow Cytometry Network (RECYF). We thank C. Klein, K. Garbin and E. Devevre for their support with the imaging. The Nanostring analysis of the NTUH core cohort was performed by the Plateforme Génomique of the Institut Curie (Paris, France). We thank D. Gentien and E. Henry for their support. We acknowledge the help of H. Yan and B. Singh.

**Author contributions** F.P., W.H.F., C.S.-F., A.d.R., T.W.-W.C., H.A.T. and A.I. designed the study and experiments. F.P., A.d.R., C.L. and Y.L. performed the bioinformatics analysis. L.L., G.L., I.N., L.-P.H., A.B., M.M. and F.P. carried out the immunohistochemistry experiments. J.C., Y.-M.J. and J.A. performed anatomo-pathology revision on the samples. E.Z.K., C.-M.S., W.-L.W. and K.M.W. performed the RNA extraction and Nanostring experiments. T.W.-W.C., A.I., M.T. and H.A.T. provided clinical guidance. T.W.-W.C., M.T., A.I., E.Z.K., A.J.L., C.L.R., M.A.B., V.B., D.R. and H.A.T. cared for the patients and provided patient materials or clinical data. F.P., W.H.F., C.S.-F., H.A.T., A.d.R., E.Z.K., C.L.R., A.J.L., T.W.-W.C., C.-M.S., J.A.W. and A.I. discussed the data and wrote the text. W.H.F., C.S.-F., A.d.R. and H.A.T. supervised the study and all authors commented on the manuscript and approved the submission.

**Competing interests** W.H.F. is a consultant for AstraZeneca, Novartis, Servier and Pierre Fabre. A.I. serves in the advisory board of Bayer, Daiichi, Epizyme, Lilly, Novartis, Roche and Springworks, and received research funding from AstraZeneca, Bayer, Chugai, Merck, MSD, Novartis and Pharmamar. J.A. is a consultant for AstraZeneca, Bayer, BMS, MSD and Roche and received research funding from MSD, Pfizer and Pierre Fabre. J.A.W. participated on advisory boards for Merck, BMS, Novartis, AstraZeneca, Roche Genentech and Illumina. M.A.B. is a consultant for EMD Serono, Immune Design, Eisai. H.A.T. serves on advisory boards and receives consulting fees from BMS, Merck and Genentech, and received research funding from BMS, Merck, Celgene, GSK, and Genentech. T.W.-W.C. participated in advisory boards for Eisai and Lilly and received research funds from Eisai. C.L.R. received research funding from BMS. The other authors declare no conflict of interest.

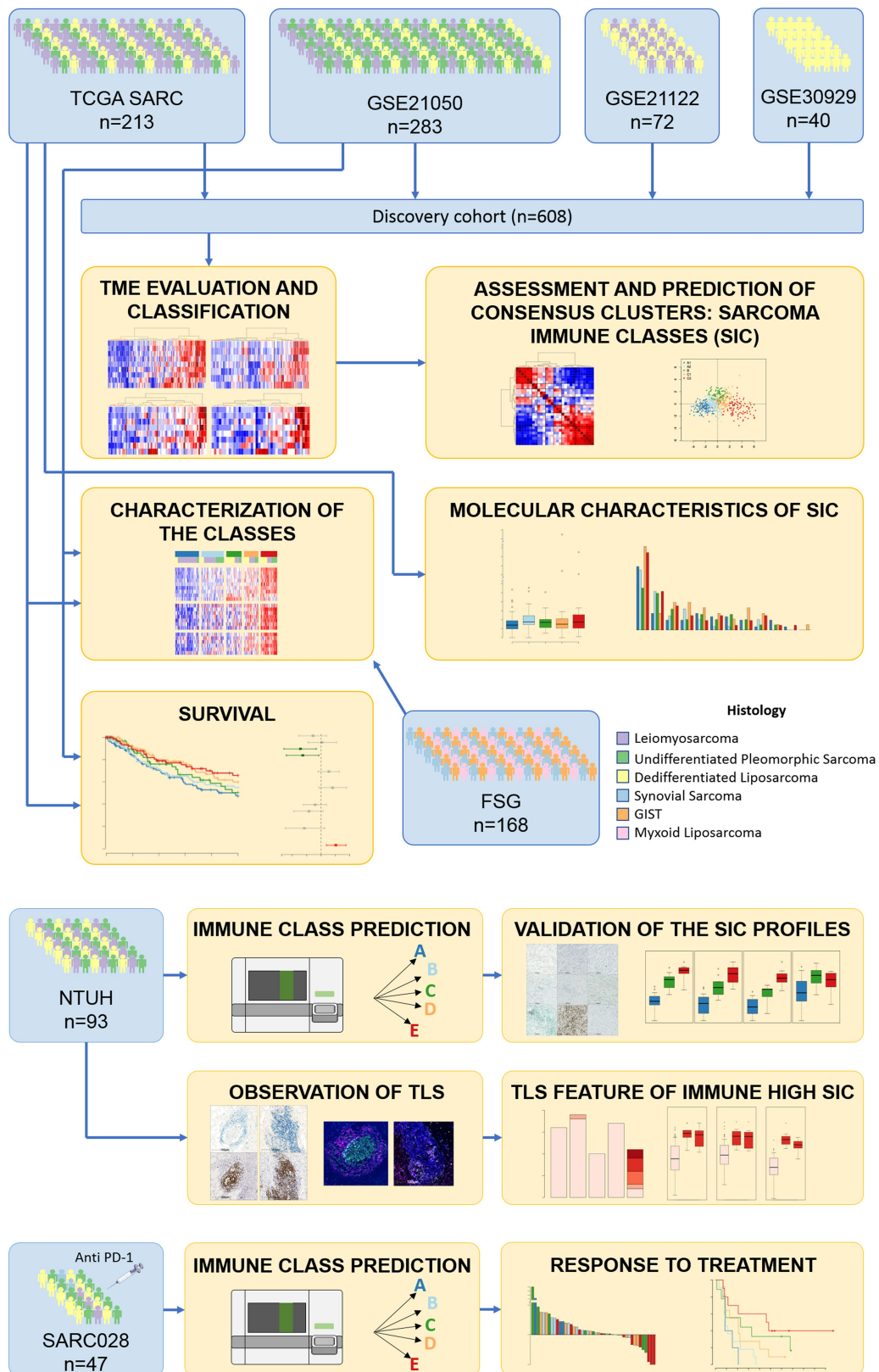
#### Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41586-019-1906-8>.

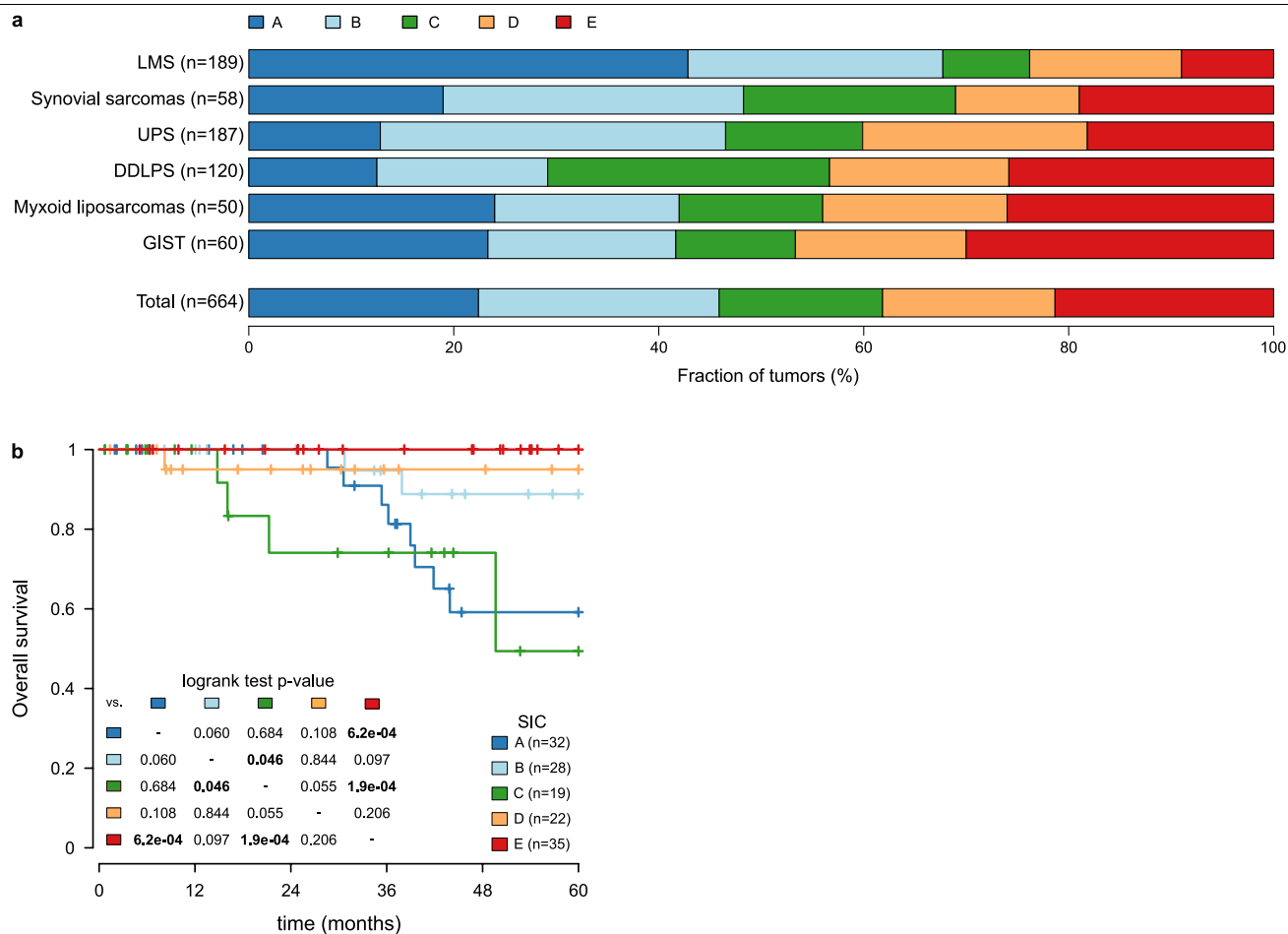
**Correspondence and requests for materials** should be addressed to H.A.T. or W.H.F.

**Peer review information** *Nature* thanks Naiyer Rizvi and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>.

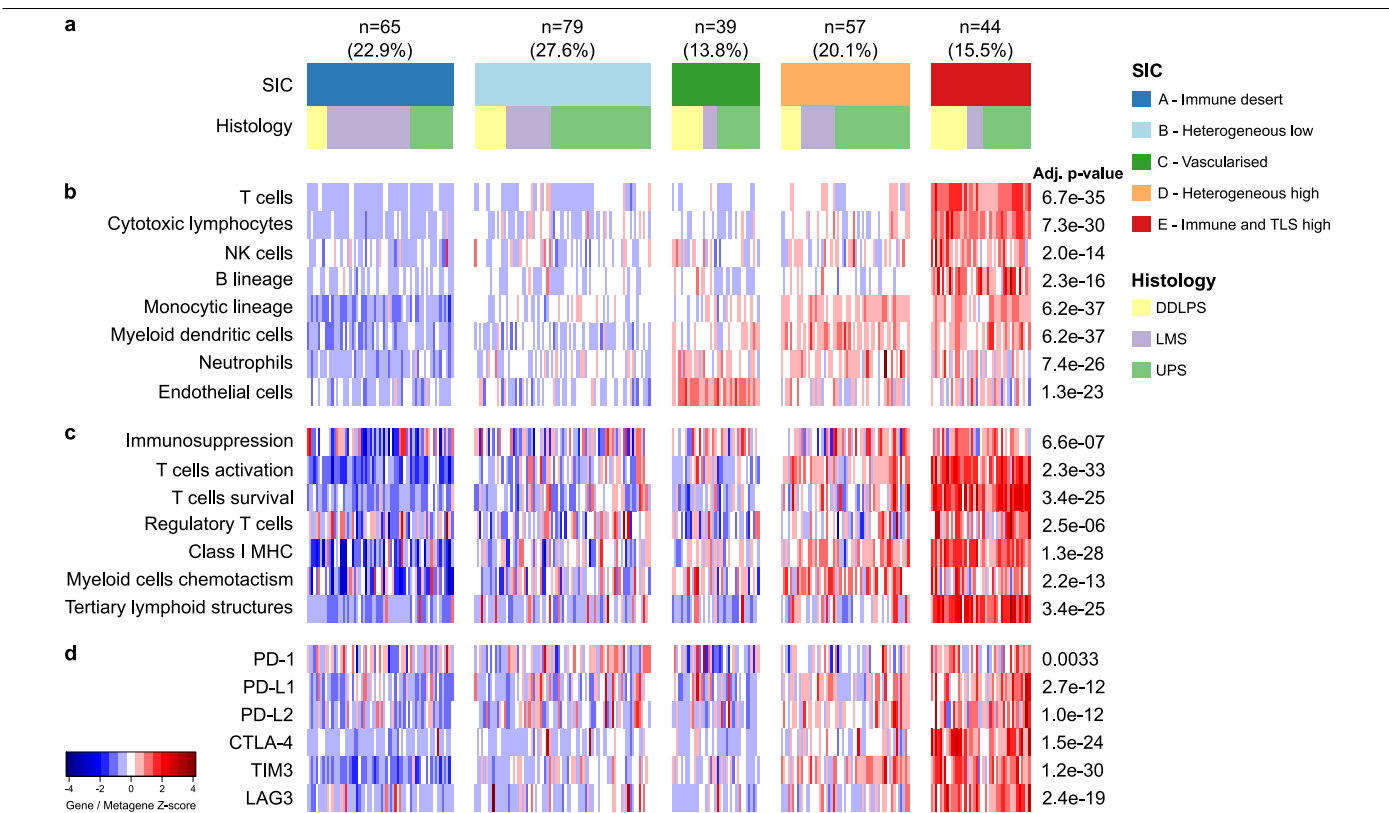


**Extended Data Fig. 1 | Diagram of analytic workflow.** The drawing of the syringe in the bottom left corner originates from Servier Medical Art (<https://smart.servier.com>), and is distributed under a CC-BY 3.0 Attribution license (<https://creativecommons.org/licenses/by/3.0/>).



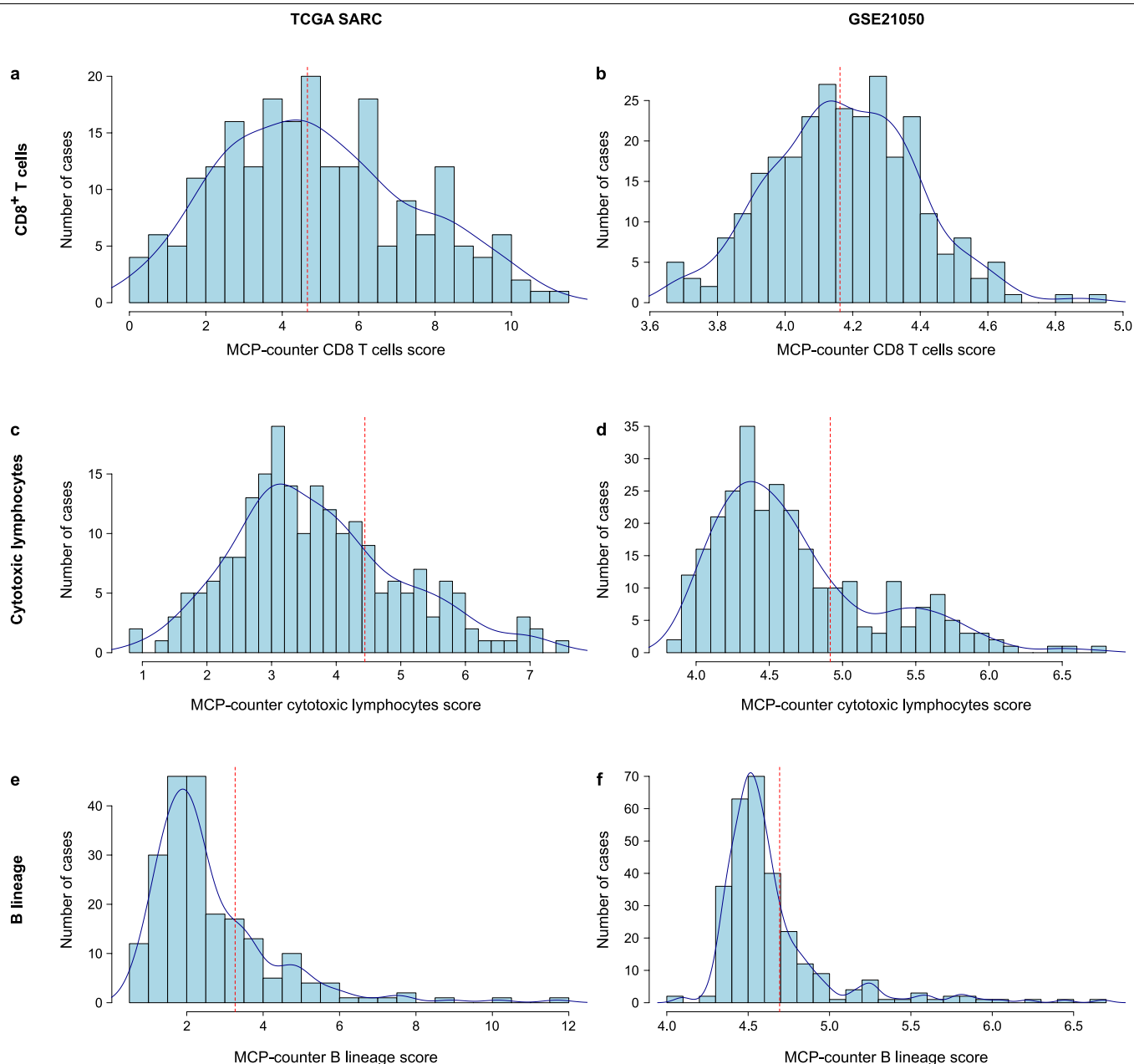
**Extended Data Fig. 2 | SICs in various STS histologies. a,** Repartition of the SICs in various histologies of TCGA SARC and GSE21050 (LMS, UPS and DDLPS), and FSG cohort (synovial sarcoma, myxoid liposarcoma, gastrointestinal stromal tumour (GIST)). **b,** Survival of patients from the FSG cohort ( $n=136$ )

according to SIC classification. Patients with synovial sarcoma, myxoid liposarcoma and gastrointestinal stromal tumour were pooled. Analysis was performed with Kaplan–Meier estimates and two-sided log-rank tests.



**Extended Data Fig. 3 | The SICs exhibit strongly different TMEs.** This figure refers to the GSE21050 cohort ( $n = 283$ ). **a**, Composition of the GSE21050 cohort by SIC, histology and site of disease. **b**, Composition of the TME by SIC as defined by the MCP-counter Z-scores. **c**, Expression of gene signatures related to the functional orientation of the immune TME by SIC. **d**, Expression

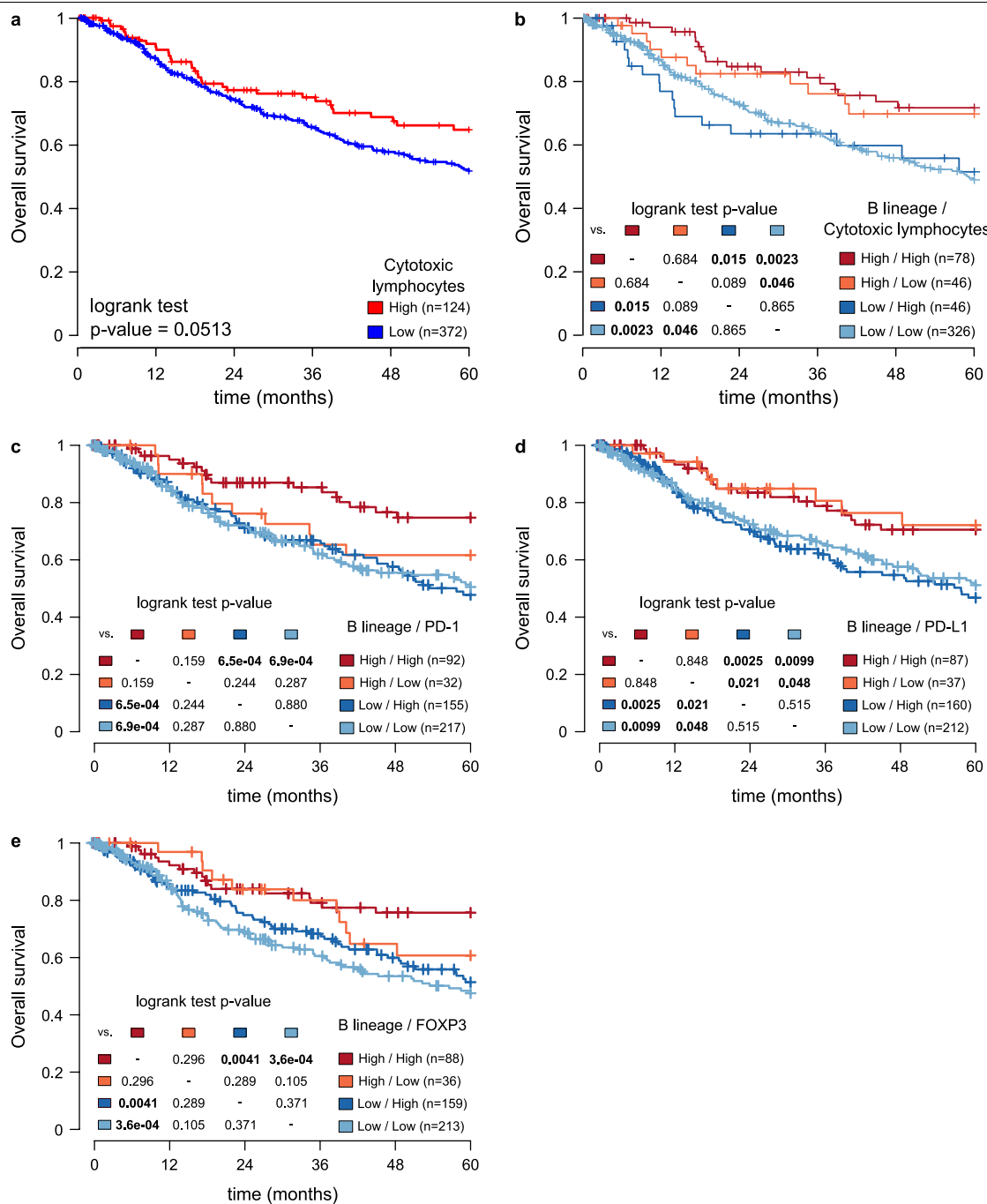
of genes related to immune checkpoints by SIC. Adjusted  $P$  values are obtained from Benjamini-Hochberg correction of two-sided Kruskal-Wallis test  $P$  values. These observations stand for cohorts GSE21122 and GSE30929 (not shown).



**Extended Data Fig. 4 | Distribution of MCP-counter scores.** **a–e**, MCP-counter scores in TCGA SARC ( $n = 213$ ) (**a, c, e**) and GSE21050 ( $n = 283$ ) (**b, d, e**), for CD8<sup>+</sup> T cells (**a, b**), cytotoxic lymphocytes (**c, d**) and B lineage cells (**e, f**). The blue line indicates the density curve. The red dotted line indicates the cut-off chosen to segregate high or low values, set at the median for CD8<sup>+</sup> T cells and at the third

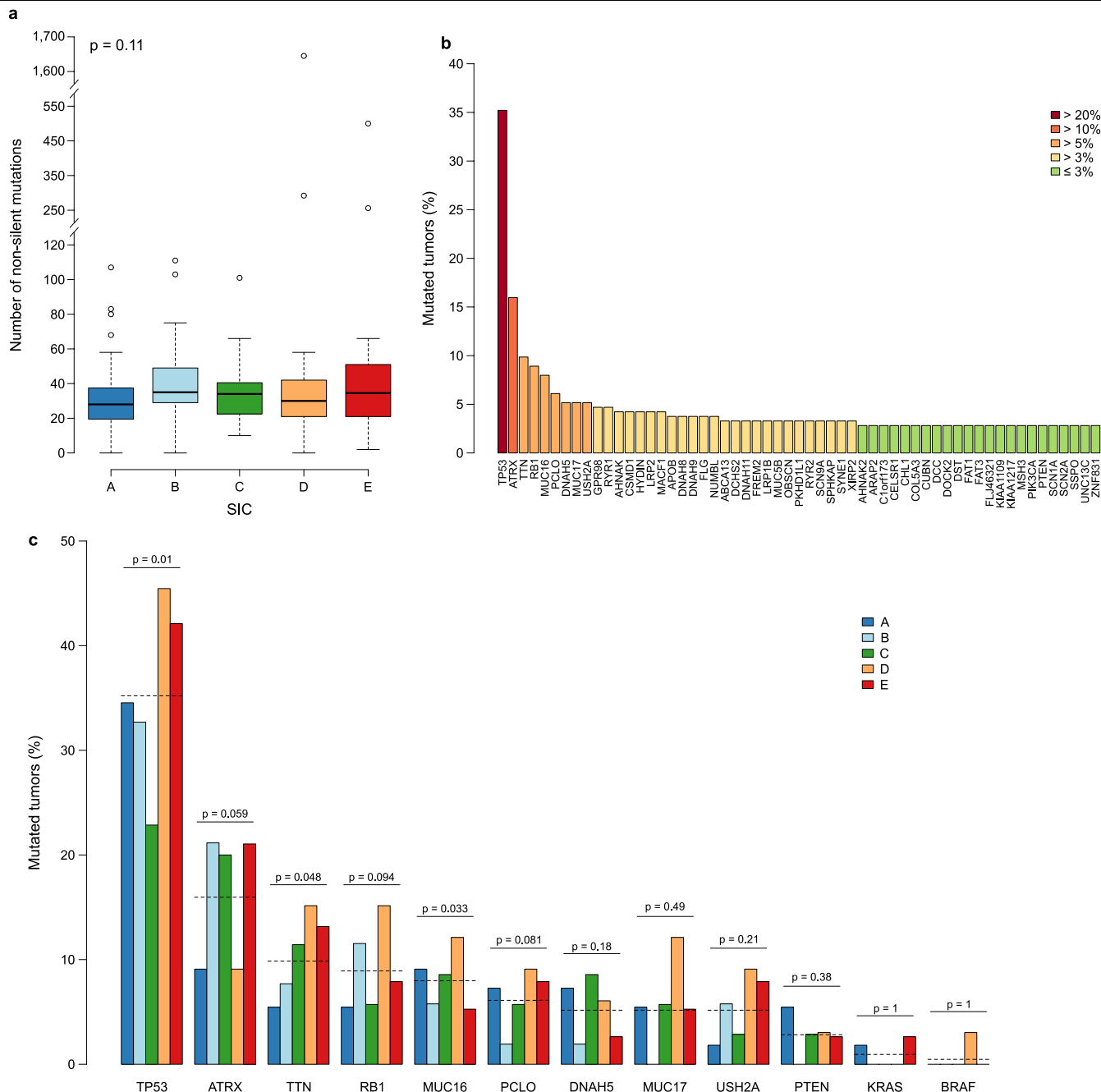
quartile for cytotoxic lymphocytes and B lineage, in each cohort. These values were chosen because the CD8 T cells scores present a normal distribution, whereas the cytotoxic lymphocytes and B lineage scores distribution exhibit a long right tail.





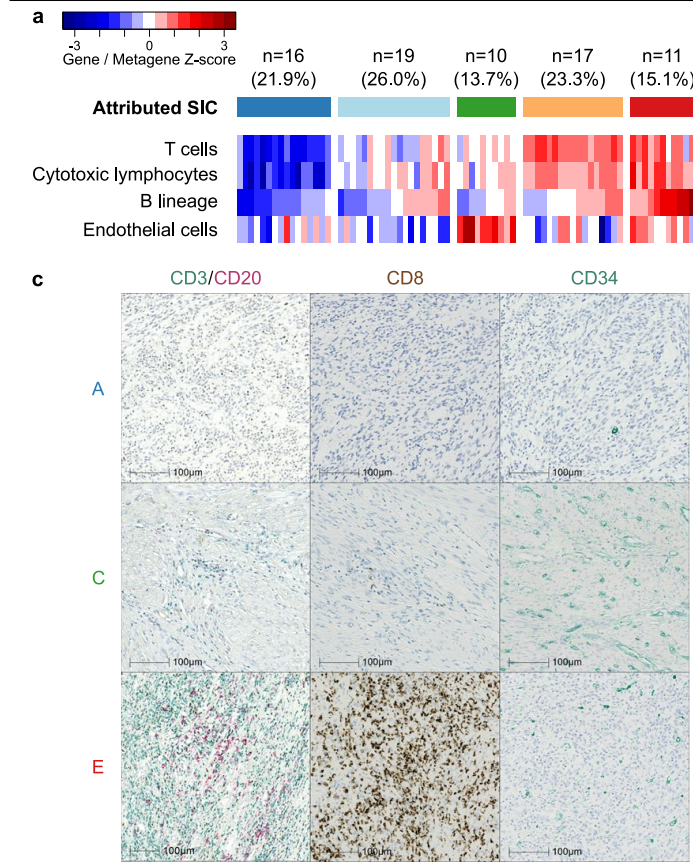
**Extended Data Fig. 5 | B cell infiltration of STS is the key factor associated with overall survival.** This figure refers to TCGA SARC and GSE21050 pooled cohorts ( $n = 496$ ). **a**, Overall survival of patients with STS according to MCP-counter scores for cytotoxic lymphocytes. **b**, Overall survival of patients based on the infiltration level of their tumours by B lineage cells and cytotoxic lymphocytes. **c–e**, Overall survival of patients based on degree of tumour

infiltration by B lineage cells and expression of *PDCD1* (**c**), *CD274* (**d**) and *FOXP3* (**e**). The analyses were performed with the Kaplan–Meier estimates and two-sided log-rank tests. Tumours were considered high for expression of *PDCD1*, *CD274* and *FOXP3* if their expression was above median, and high for B lineage and cytotoxic lymphocytes if the MCP-counter score was above the third quartile.

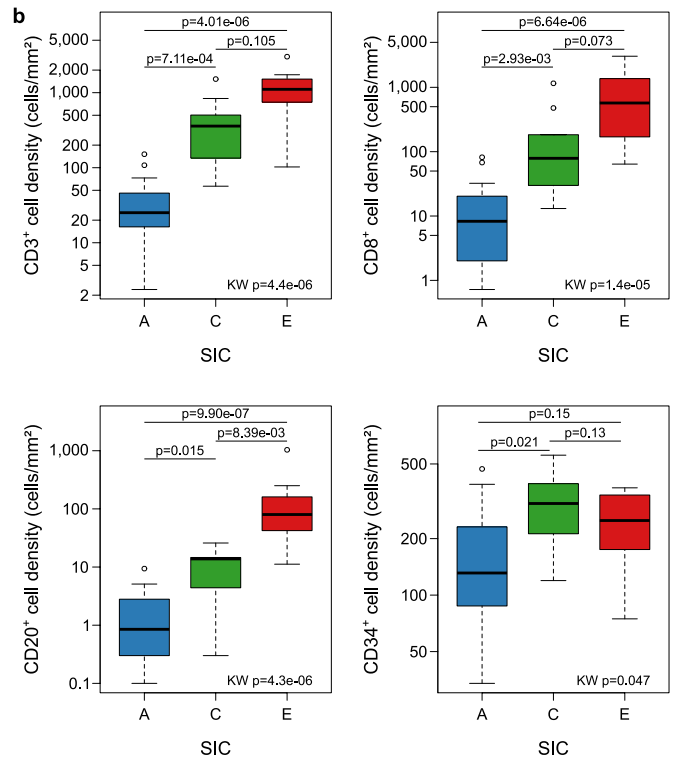


**Extended Data Fig. 6 | The mutational landscape of STS tumours does not vary significantly between SICs.** This figure refers to the TCGA SARC cohort ( $n = 213$ ). **a**, Mutational burden according to the SIC of the tumours, expressed in number of non-silent mutations.  $P$  value was computed with a Kruskal-Wallis test. Box plots as in Fig. 3d. **b**, Mutation frequency of all genes that are mutated in greater than 2.5% of tumours. **c**, Mutation frequency for genes that are mutated in more than 5% of tumours, according to SICs in the TCGA SARC

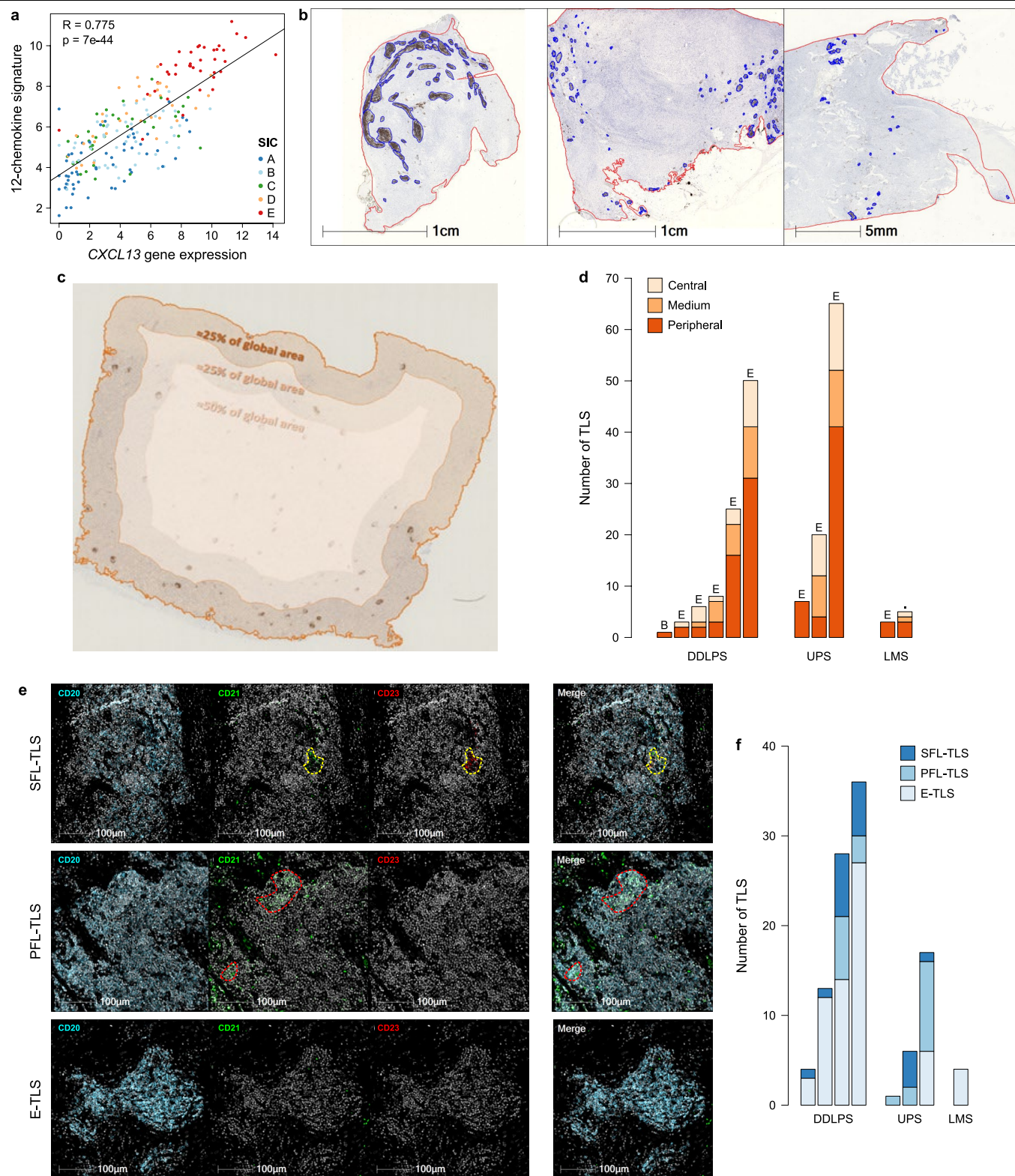
cohort. The dashed lines indicate the overall mutation frequency.  $P$  values were obtained through one-sample two-sided  $t$ -tests, corrected for multiple testing with the Bonferroni method. This was applied only to samples that had mutations on the considered genes (*TP53*:  $n = 75$ ; *ATRX*:  $n = 34$ ; *TTN*:  $n = 21$ ; *RB1*:  $n = 19$ ; *MUC16*,  $n = 17$ ; *PCLO*,  $n = 13$ ; *DNAH5*, *MUC17* and *USH2A*:  $n = 11$ , *PTEN*,  $n = 6$ ; *KRAS*,  $n = 2$ ; *BRAF*,  $n = 1$ ).



**Extended Data Fig. 7 | Validation of SIC profiles by immunohistochemistry.** This figure refers to the NTUH cohort. **a**, SIC attribution as defined by gene expression using the MCP-counter Z-scores in 73 cases. **b**, Cell density counts showing the differences in TME composition according to SIC identification of the 73 cases (SIC A:  $n=16$ ; SIC C:  $n=10$ ; SIC E:  $n=11$ ). P-values are determined by two-sided Kruskal–Wallis (KW) tests. Pairwise comparisons are derived from



the Dunn test. Box plots are as in Fig. 3d. **c**, Representative images of CD3 (green), CD20 (pink), CD8 (brown) and CD34 (green) expression by immunohistochemistry of SIC A, C and E tumours. The same area of the tumour is represented (0.05 mm<sup>2</sup>) in each image. Similar results were observed on the other tumours from the same SICs (SIC A: *n* = 16; SIC C: *n* = 10; SIC E: *n* = 11).

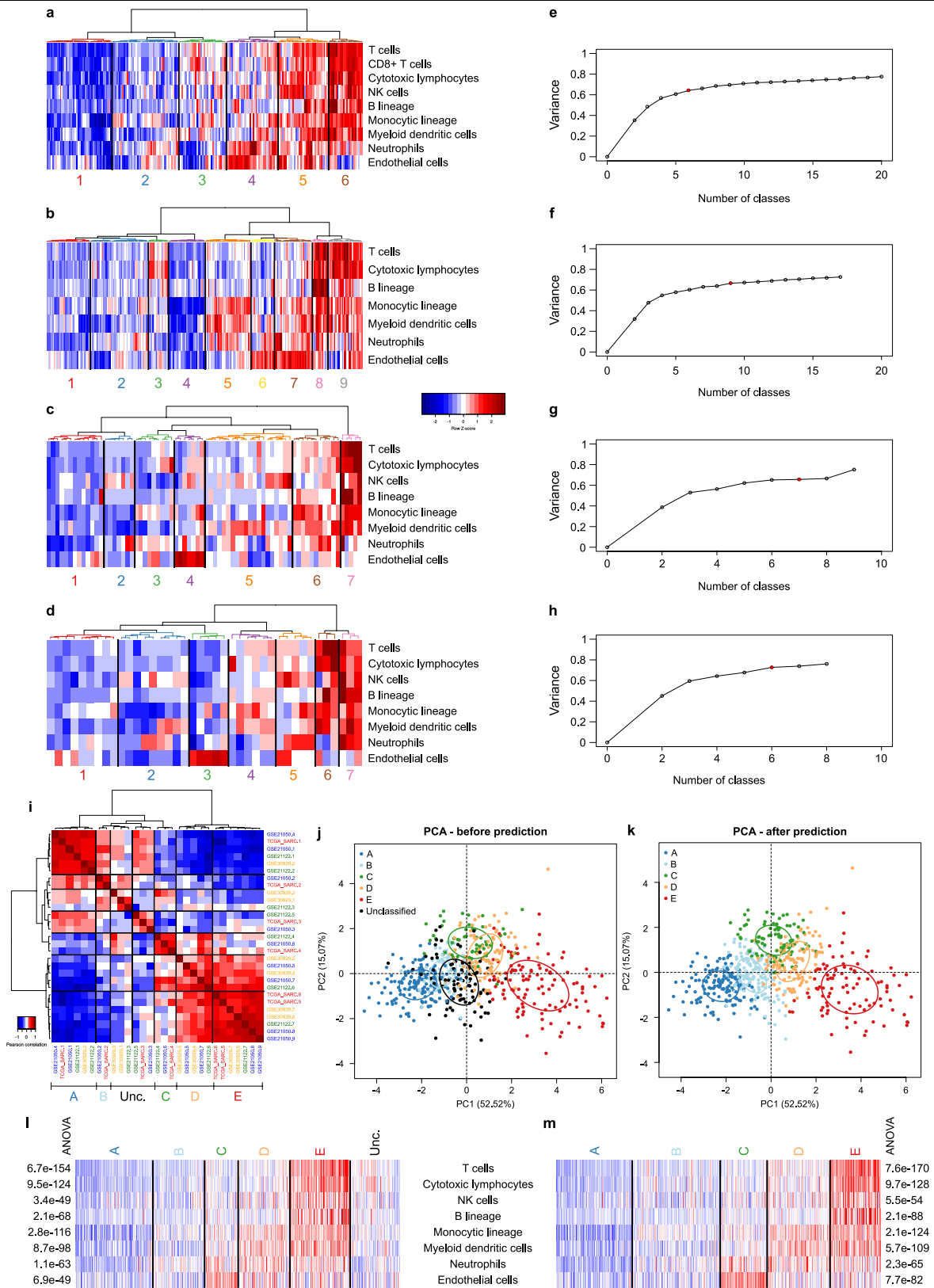


**Extended Data Fig. 8** | See next page for caption.

**Extended Data Fig. 8 | Location and maturation of TLSs.** **a**, Pearson correlation between the expression of CXCL13 and the 12-chemokine signature of TLS in TCGA SARC cohort ( $n = 213$ ). Samples are coloured according to SICs. **b**, Intratumoural location of TLSs in three different examples from the NTUH cohort—DDLPS, UPS and LMS, respectively. TLSs are observed by the presence of CD20<sup>+</sup> B cells aggregates (brown, surrounded by blue shapes). The red line delineates the tumoral zone. Similar findings were observed on the 11 tumours with TLS. **c**, Definition of peripheral, medium and central zones, accounting for 25%, 25% and 50% of the total tumour area, respectively. **d**, Distribution of TLSs in the various zones. Each bar represents one tumour. The letters above bars indicate the SIC of the tumour when the sample passed quality control of Nanostring nCounter hybridization. Dots indicate tumours in which SIC could not be determined because of RNA quality control. Similar images were observed for 66 E-TLS, 23 PFL-TLS and 20 SFL-TLS. **e**, Illustration of diverse

degrees of TLS maturation in STS tumours. Consistent with maturation events occurring in secondary lymphoid organs, three maturation steps have been described for TLS: E-TLS (bottom), PFL-TLS (middle) and SFL-TLS (top), which differ in the presence of follicular dendritic cells (FDC) and their markers. E-TLS contain aggregates of CD20<sup>+</sup> B cells and CD3<sup>+</sup> T cells without FDC, PFL-TLS contain CD21<sup>+</sup> FDC (red dotted zones) and SFL-TLS contain a germinal centre, notably visible through the presence of CD21<sup>+</sup>CD23<sup>+</sup> follicular dendritic cells (yellow dotted zone). DAPI staining is shown in white. DAPI-negative green dots correspond to fluorescent erythrocytes. **f**, Distribution of TLS maturation steps in a subset of tumours. Each bar represents one tumour. Differences between the number of TLSs observed here and in other figures can be explained by use of non-consecutive slides or a different tumour block for some samples.





Extended Data Fig. 9 | See next page for caption.

**Extended Data Fig. 9 | Pan-cohort immune classification.** This figure refers to the four discovery cohorts: TCGA SARC ( $n = 213$ ), GSE21050 ( $n = 283$ ), GSE21122 ( $n = 72$ ) and GSE30929 ( $n = 40$ ). **a–d**, Heat map and unsupervised hierarchical clustering of the MCP-counter scores describing the tumour microenvironment. Each of the population is represented by the Z-scores of the signature. **a**, TCGA SARC. **b**, GSE21050. **c**, GSE21122. **d**, GSE30929. **e–h**, Evolution of the variance explained by the clusters as a function of the number of clusters. Red dots indicate the number of clusters that was retained in this study. Each graph corresponds to the heat map on its left. **i**, Heat map of the Pearson correlation of centroids from each SIC class of discovery cohorts (TCGA SARC, GSE21050, GSE21122 and GSE30929,  $n = 608$ ), with five immune classes and two groups of unclassified samples. **j**, **k**, Principal component analysis of samples from the four discovery cohorts ( $n = 608$ ), based on their normalized and merged MCP-counter scores. **j** is coloured according to the original classes, **k** is coloured according to the predicted immune classes, showing a heightened homogeneity within each SIC class. **l**, **m**, Composition of the TME with classes defined as in **j** and **k** for the four discovery cohorts ( $n = 608$ ), expressed in cohort-specific row Z-scores.

**Extended Data Table 1 | Clinicopathological composition of the cohorts included in this study**

	Cohort	TCGA SARC	GSE21050	GSE21122	GSE30929	FSG	NTUH	SARC028
	n	213	283	72	40	168	93 (SIC: 73)	47
Age (median, range)		63 (33-90)	63 (15-92)	ND	ND	36 (1-83)	58 (9-94)	57 (25-83)
Gender (n, %)								
	Male	98 (46%)	131 (48.7%)	ND	ND	99 (58.9%)	37 (39.8%)	34 (72.3%)
	Female	114 (54%)	138 (51.3%)	ND	ND	69 (41.1%)	56 (60.2%)	13 (27.7%)
STS histology (n,%)								
	DDLPS	58 (27.2%)	62 (21.9%)	46 (63.9%)	40 (100%)	0 (0%)	30 (32.3%)	19 (40.4%)
	LMS	104 (48.8%)	85 (30%)	26 (36.1%)	0 (0%)	0 (0%)	31 (33.3 %)	6 (12.8%)
	UPS	51 (23.9%)	136 (48.1%)	0 (0%)	0 (0%)	0 (0%)	32 (34.4%)	19 (40.4%)
	Synovial sarcoma	0 (0%)	0 (0%)	0 (0%)	0 (0%)	58 (34.5%)	0 (0%)	3 (6.4%)
	Myxoid liposarcoma	0 (0%)	0 (0%)	0 (0%)	0 (0%)	50 (29.8%)	0 (0%)	0 (0%)
	GIST	0 (0%)	0 (0%)	0 (0%)	0 (0%)	60 (35.7%)	0 (0%)	0 (0%)
SIC (n,%)								
	A	55 (25.8%)	65 (23%)	14 (19.4%)	8 (20%)	37 (22%)	16 (21.9%)	5 (10.6%)
	B	52 (24.4%)	78 (27.6%)	22 (30.6%)	13 (32.5%)	37 (22%)	19 (26%)	11 (23.4%)
	C	35 (16.4%)	39 (13.9%)	8 (11.1%)	6 (15%)	26 (15.5%)	10 (13.7%)	9 (19.1%)
	D	33 (15.5%)	57 (20.1%)	21 (29.2%)	7 (17.5%)	26 (15.5%)	17 (23.3%)	12 (25.5%)
	E	38 (17.8%)	44 (15.5%)	7 (9.7%)	6 (15%)	42 (25%)	11 (15.1%)	10 (21.3%)

For cohort GSE21050, sex information could not be retrieved for 14 patients. For cohort NTUH, SIC could be determined for 73 patients only. NA, not available.

Extended Data Table 2 | Antibodies used for immunohistochemistry and immunofluorescence

Antibody	References	Species	Clone	Source	Concentration (ug/mL)	Antigen Retrieval	Secondary Antibody	Tertiary Reagents
Immunohistochemistry								
CD3	790-4341	Rabbit	2GV6	Roche	0.4	pH9	Alkaline Phosphatase Goat Anti-Rabbit IgG	Alkaline Phosphatase Blue Substrate
CD8	M7103	Mouse IgG1	C8/144B	Agilent	1.57	pH9	EnVision+ System-HRP, Labelled Polymer (Mouse)	DAB
CD20	M0755	Mouse IgG2a	L26	Agilent	0.6	pH6	EnVision+ System-HRP, Labelled Polymer (Mouse)	DAB
							Polyview Plus AP (anti-mouse) reagent	HighDef red IHC chromogen (AP)
CD21	M0784	Mouse IgG1	1F8	Agilent	6.6	pH6	EnVision+ System-HRP, Labelled Polymer (Mouse)	DAB
CD23	ab16702	Rabbit	SP23	Abcam	0.5	pH9	EnVision+ System-HRP, Labelled Polymer (Rabbit)	Permanent HRP Green
CD34	M7165	Mouse IgG1	Qbend-10	Agilent	0.48	pH6	EnVision+ System-HRP, Labelled Polymer (Mouse)	Permanent HRP Green
DC-Lamp	DDX0191	Rat IgG2a	1010E1.01	Dendritics	5	pH9	Biotin Donkey Anti-Rat IgG	AEC Peroxidase (HRP) Substrate
PD-L1	13684	Rabbit	E1L3N	Cell Signaling	1.25	pH9	EnVision+ System-HRP, Labelled Polymer (Rabbit)	DAB
PNAd	553863	Rat IgM	MECA-79	BD	5	pH6	ImmPRESS HRP Anti-Rat (Peroxidase) Polymer Detection Kit	DAB
Immunofluorescence								
CD3	790-4341	Rabbit	2GV6	Roche	0.4	pH9	EnVision+ System-HRP, Labelled Polymer (Rabbit)	Alexa Fluor™ 647 Tyramide Reagent
CD4	AC-0173	Rabbit	EP204	Epitomics	0.17	pH9	Cy3 Donkey Anti-Rabbit IgG	NA
CD20	M0755	Mouse IgG2a	L26	Agilent	0.6	pH6	EnVision+ System-HRP, Labelled Polymer (Mouse)	Alexa Fluor™ 594 Tyramide Reagent
					1.2	pH 6	Cy5 Goat Anti-Mouse IgG2a	NA
CD21	M0784	Mouse IgG1	1F8	Agilent	7.5	pH 6	Cy3 Goat Anti-Mouse IgG1	NA
CD23	ab16702	Rabbit	SP23	Abcam	2	pH 6	FITC Donkey Anti-Rabbit IgG	NA
CXCR5	MAB190-100	Mouse IgG2b	51505	R&D	5	pH9	FITC Rat Anti-Mouse IgG2b	NA
PD-1	NA	Mouse IgG2a	EH33	CoStim	3	pH6	EnVision+ System-HRP, Labelled Polymer (Mouse)	Alexa Fluor™ 488 Tyramide Reagent
					6	pH9	Cy5 Goat Anti-Mouse IgG2a	NA

## Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- ☐ ☒ The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- ☐ ☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- ☐ ☒ The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- ☐ ☒ A description of all covariates tested
- ☐ ☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- ☐ ☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- ☐ ☒ For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give  $P$  values as exact values whenever suitable.*
- ☒ ☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- ☒ ☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- ☐ ☒ Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

### Software and code

Policy information about [availability of computer code](#)

Data collection

Immunohistochemistry images were analysed with HALO 10 software (IndicaLab). Immunofluorescence data were obtained with AxioScan (Zeiss)

Data analysis

Data was analysed with R software (version 3.4.4) and packages gplots, survival and FactoMineR. Custom code was produced in R for the analysis.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

### Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The transcriptomic datasets analysed in this study can be accessed on the GDC Portal (TCGA SARC) and the Gene Expression Omnibus repository (accession numbers GSE21050, GSE21122, GSE30929). Immunohistochemistry, gene expression and clinical-related to NTUH cohorts (Fig. 3, Extended Data Figs. 7 and 8) are available from the corresponding author on reasonable request. The data that support the findings related to Fig. 4 are available from SARC but restrictions apply to the availability of these data, which were used under license for the study. Data are however available upon reasonable request to HAT (HTawbi@mdanderson.org) and with permission of SARC. All code used in this study is available from the authors upon reasonable request.



## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	TCGA SARC: n=213, GSE21050: n=283, GSE21122: n=72, GSE30929: n=40, FSG: n=168, NTUH n=93, SARC028: n=47. Total: n=916.
Data exclusions	20 tumours from the NTUH cohort were excluded from gene expression (end SIC) analysis due to low quality of the extracted RNA.
Replication	No replication was done, but validation cohorts were analysed.
Randomization	Randomization is only relevant to the SARC028 cohort, which was previously published.
Blinding	All image and data analysis were performed blindly, independently of sample knowledge.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input type="checkbox"/>	<input checked="" type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data

### Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

## Antibodies

Antibodies used	CD3: 2GV6, Roche ; DC-Lamp: 1010E1.01, Dendritics ; CD20: L26, AGilent ; CD8: C8/144B, Agilent ; CD21 : 1F8, Agilent ; CD23 : SP23, Abcam; CD34: Qbend-10, Agilent ; PD-L1: E1L3N, Cell Signaling ; PD-1: EH33, CoStim Pharmaceuticals ; .
Validation	The specificity of anti-CD3, anti-CD4, anti-CD8, anti-CD20, anti-CD21, anti-CD23, anti-CD34, anti-CXCR5 and anti-DC-Lamp antibodies, and MECA-79 (PNAd) was validated on FFPE tonsil sections as positive control. For anti-CD20, certified manufacturing facilities from the company guarantee full quality control including western blot and studies using COS-1 cells transfected with cDNA encoding the CD20 molecule indicate that the antibody labels an intracytoplasmic epitope localized on the CD20 molecule. For anti-CD8, certified manufacturing facilities from the company guarantee full quality control including western blot and indicate that the antibody recognizes the cd8alpha chain. For anti-CD34, certified manufacturing facilities from the company guarantee full quality control. For anti-CD21, certified manufacturing facilities from the company guarantee full quality control including western blotting of the immunogen, and that the antibody labels cells or cell lines known to express CD21 (Raji, NC 37, tonsil cells), whereas no labeling is observed in the CD21-negative Jurkat cells (T-cell line) and human erythrocytes. For anti-CD23, certified manufacturing facilities from the company guarantee full quality control including western blotting, IHC on human tonsils and flow cytometry on Raji cells. For anti-CXCR5, certified manufacturing facilities from the company guarantee full quality control using human CXCR5 transfectants by flow cytometry and lack of cross reactivity with human CXCR2, CXCR3, or CXCR4 transfectants. For PNAd, certified manufacturing facilities from the company guarantee full quality control including western blotting, IHC and flow cytometry. For anti-PD-L1, specificity was validated by the company using immunohistochemical analysis of paraffin-embedded human placenta using PD-L1 (E1L3N®) XP® Rabbit mAb in the presence of control peptide or antigen-specific peptide. Specificity was verified by using FPE sections from placenta as positive control and cerebral cortex tissue as negative control. Anti-PD-1 (Freeman GJ and col.) was obtained from CoStim Pharmaceuticals and validated as described in Fig. S1 of Giraldo et al., Clinical Cancer Research, 2015. Tonsil, placenta and cerebral cortex slides were obtained from Geneticist Inc.

# Human research participants

Policy information about [studies involving human research participants](#)

Population characteristics	All available characteristics are reported in Extended Data Table 1.
Recruitment	Patients were recruited prior to the study and were not selected on specific criteria other than their pathology.
Ethics oversight	The research was approved by the Research Ethics Committee of NTUH (201605061RINA).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

# Tertiary lymphoid structures improve immunotherapy and survival in melanoma

<https://doi.org/10.1038/s41586-019-1914-8>

Received: 5 February 2019

Accepted: 4 December 2019

Published online: 15 January 2020

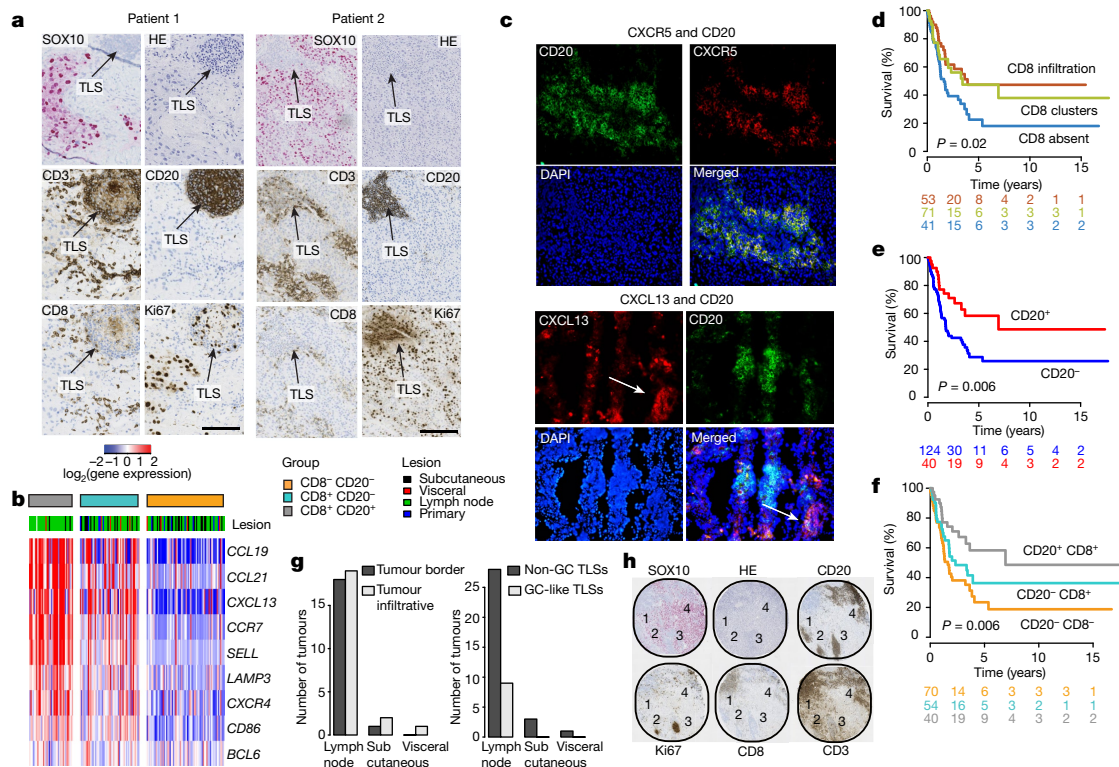
Rita Cabrita<sup>1,12</sup>, Martin Lauss<sup>1,12</sup>, Adriana Sanna<sup>1</sup>, Marco Donia<sup>2</sup>, Mathilde Skaarup Larsen<sup>3</sup>, Shamik Mitra<sup>1</sup>, Iva Johansson<sup>1</sup>, Bengt Phung<sup>1</sup>, Katja Harbst<sup>1</sup>, Johan Vallon-Christersson<sup>1</sup>, Alison van Schoiack<sup>4</sup>, Kristina Lövgren<sup>1</sup>, Sarah Warren<sup>4</sup>, Karin Jirstrom<sup>1</sup>, Håkan Olsson<sup>1</sup>, Kristian Pietras<sup>5</sup>, Christian Ingvar<sup>6</sup>, Karolin Isaksson<sup>6</sup>, Dirk Schadendorf<sup>7</sup>, Henrik Schmidt<sup>8</sup>, Lars Bastholt<sup>9</sup>, Ana Carneiro<sup>1,10</sup>, Jennifer A. Wargo<sup>11</sup>, Inge Marie Svane<sup>2</sup> & Göran Jönsson<sup>1\*</sup>

Checkpoint blockade therapies that reactivate tumour-associated T cells can induce durable tumour control and result in the long-term survival of patients with advanced cancers<sup>1</sup>. Current predictive biomarkers for therapy response include high levels of intratumour immunological activity, a high tumour mutational burden and specific characteristics of the gut microbiota<sup>2,3</sup>. Although the role of T cells in antitumour responses has thoroughly been studied, other immune cells remain insufficiently explored. Here we use clinical samples of metastatic melanomas to investigate the role of B cells in antitumour responses, and find that the co-occurrence of tumour-associated CD8<sup>+</sup> T cells and CD20<sup>+</sup> B cells is associated with improved survival, independently of other clinical variables. Immunofluorescence staining of CXCR5 and CXCL13 in combination with CD20 reveals the formation of tertiary lymphoid structures in these CD8<sup>+</sup>CD20<sup>+</sup> tumours. We derived a gene signature associated with tertiary lymphoid structures, which predicted clinical outcomes in cohorts of patients treated with immune checkpoint blockade. Furthermore, B-cell-rich tumours were accompanied by increased levels of TCF7<sup>+</sup> naive and/or memory T cells. This was corroborated by digital spatial-profiling data, in which T cells in tumours without tertiary lymphoid structures had a dysfunctional molecular phenotype. Our results indicate that tertiary lymphoid structures have a key role in the immune microenvironment in melanoma, by conferring distinct T cell phenotypes. Therapeutic strategies to induce the formation of tertiary lymphoid structures should be explored to improve responses to cancer immunotherapy.

In addition to T cells, the main component of the adaptive immune system consists of B cells. B cells localized in so-called tertiary lymphoid structures (TLSs)—which have been identified in several types of cancer, including melanoma<sup>4–6</sup>—may improve antigen presentation, increase cytokine-mediated signalling, release tumour-specific antibodies, are associated with improved prognosis<sup>7</sup> and, to some extent, with clinical responses to CTLA4<sup>5</sup>. Additional evidence on the importance of TLSs in the tumour immune microenvironment is provided in the accompanying Articles<sup>8,9</sup>. In our analysis of the immune microenvironment of melanoma tumours, we found infiltration of CD8<sup>+</sup> T cells in 33% of cases; 25% of the tumours had CD8<sup>+</sup> T cells localized in clusters, and 42% were devoid of CD8<sup>+</sup> T cells (Extended Data Table 1). By contrast, we found CD20<sup>+</sup> B cell clusters in 25% of the cases and such clusters consisted of both Ki67<sup>+</sup> and Ki67<sup>−</sup> B cells (Fig. 1a), which suggests that some B cells are activated and proliferating<sup>10</sup>. Notably, CD20<sup>+</sup> B cell clusters

were in all cases surrounded mainly by CD4<sup>+</sup> T cells, which indicates formation of TLSs (Extended Data Fig. 1a). We then analysed whether these CD20<sup>+</sup> B cell clusters have similarities to bona fide TLSs. Known molecular markers of TLS formation include increased expression of CXCL13, CXCR5 and DC-LAMP<sup>11</sup>. These markers were all upregulated in transcriptomic data from matched tumour tissue (Fig. 1b). Moreover, immunofluorescence staining of two known TLS markers (CXCR5 and CXCL13), in combination with CD20, supported the notion that these CD20<sup>+</sup> B cell clusters have molecular properties that have been described as necessary for TLS formation<sup>11</sup> (Fig. 1c). By contrast, CD8<sup>+</sup> T cells were localized mainly outside of such TLSs, but the presence of TLSs was in all cases coupled with tumour-associated CD8<sup>+</sup> T cells (Fig. 1a, Extended Data Table 1). The formation of TLSs may indicate that tumour antigens are recognized by the immune system. The inability of the immune system to completely eradicate the tumour would

<sup>1</sup>Department of Clinical Sciences, Division of Oncology and Pathology, Lund University Cancer Center, Lund University, Lund, Sweden. <sup>2</sup>National Center for Cancer Immune Therapy, Department of Oncology, Copenhagen University Hospital, Herlev, Denmark. <sup>3</sup>Department of Clinical Pathology, Herlev University Hospital, Herlev, Denmark. <sup>4</sup>NanoString Technologies, Seattle, WA, USA. <sup>5</sup>Division of Translational Cancer Research, Department of Laboratory Medicine, Lund University Cancer Center, Lund University, Lund, Sweden. <sup>6</sup>Department of Surgery, Skåne University Hospital, Lund, Sweden. <sup>7</sup>Department of Dermatology, University Hospital of Essen, Essen, Germany. <sup>8</sup>Department of Oncology, Århus University Hospital, Aarhus, Denmark. <sup>9</sup>Department of Oncology, Odense University Hospital, Odense, Denmark. <sup>10</sup>Department of Oncology, Skåne University Hospital, Lund, Sweden. <sup>11</sup>Department of Surgical Oncology, MD Anderson Cancer Center, Houston, TX, USA. <sup>12</sup>These authors contributed equally: Rita Cabrita, Martin Lauss. \*e-mail: goran\_b.jonsson@med.lu.se



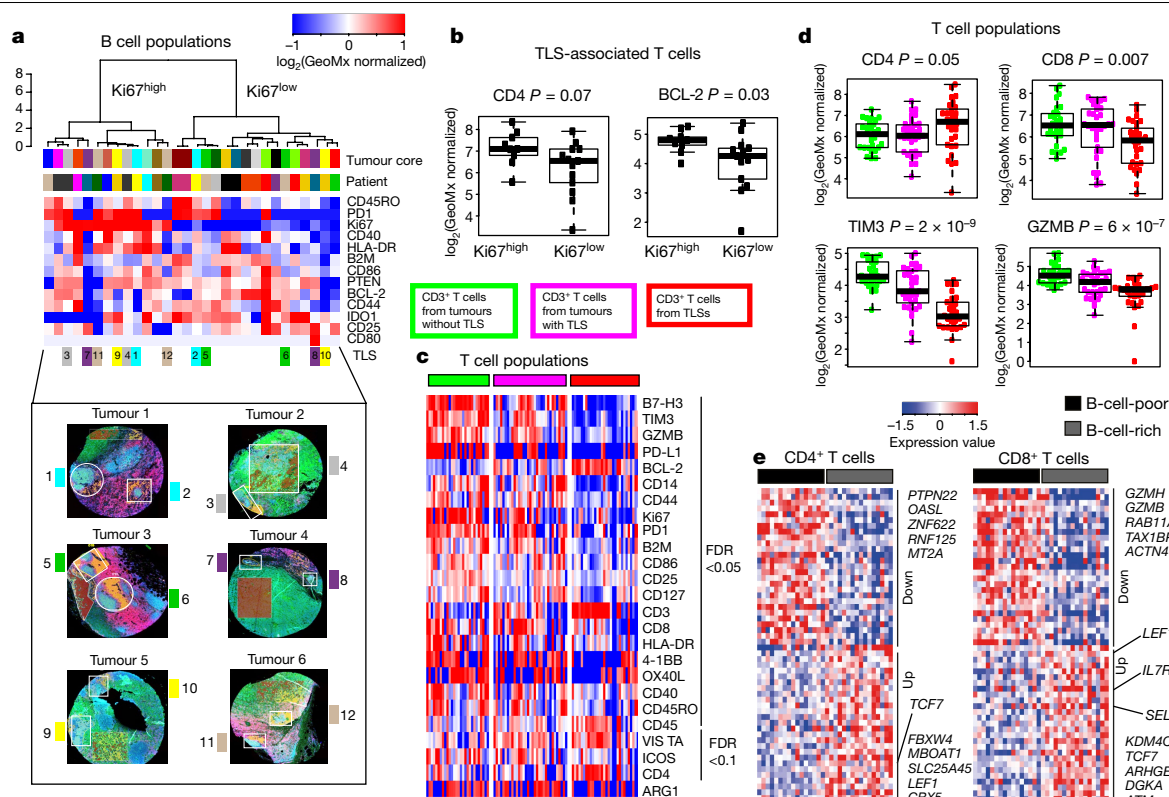
**Fig. 1 | Identification of CD20<sup>+</sup> B cell clusters in melanoma tumours.**

**a**, Representative immunostaining of CD20 (B cells), Ki67 (proliferating cells), SOX10 (melanoma cells), CD3 (T cells) and CD8 (T cells). In total, 177 melanoma specimens—including 113 lymph node metastases, 35 subcutaneous metastases, 10 visceral metastases and 15 primary tumours—were analysed. Sections were taken consecutively to spatially analyse the different immunostainings. Scale bars, 100  $\mu$ m (patient 1), 200  $\mu$ m (patient 2). HE, haematoxylin and eosin stain. **b**, Gene-expression heat map of known TLS marker genes. The gene-expression data were obtained from matched tumour tissue ( $n = 160$ ), as was used for the immunostaining. **c**, Representative immunofluorescence staining of CD20 (green) in combination with CXCR5 (red) or CXCL13 (red) in a melanoma tumour known to have TLSs, selected from the immunostaining cohort in **a**. Arrows indicate a CXCL13<sup>+</sup> cell cluster.

**d–f**, Kaplan–Meier survival analysis of the cohort stratified by CD8 (**d**), CD20 (**e**) and combining these two markers (**f**);  $n = 165$ ,  $n = 164$  and  $n = 164$  patients with available follow-up information in **d**, **e** and **f**, respectively. Cox regression analysis was used to calculate  $P$  values. Numbers below plots represent numbers of patients. **g**, TLSs were evaluated for the level of maturation using Ki67 immunostaining and spatial location. Mature germinal-centre (GC)-like structures were detected exclusively in TLSs located in lymph node metastases, and there was no difference in spatial location between lymph node metastases and others. **h**, A representative case with multiple TLSs (numbered 1–4). TLS 2 and TLS 3 show a germinal-centre-like structure within the TLS, whereas TLS 1 and TLS 4 lack these structures. The representative case was selected from  $n = 18$  investigated cases with multiple TLSs.

subsequently lead to chronic inflammation, which is characterized by infiltrating immune cells and generation of TLSs<sup>12</sup>. Importantly, because all tumours with TLSs had tumour-associated CD8<sup>+</sup> T cells (Fig. 1a, Extended Data Table 1), we hypothesized that TLSs may support the activation of CD8<sup>+</sup> T cell attack against tumour cells. Indeed, survival analysis revealed that the presence of tumour-associated CD8<sup>+</sup> T cells or TLSs was associated with improved patient outcome in uni- and multivariate analyses (Fig. 1d, e, Extended Data Table 2). The combination of both TLSs and CD8<sup>+</sup> T cells was associated with the best survival outcome, CD8<sup>+</sup> T cells alone was linked with intermediate survival, and the absence of both TLSs and CD8<sup>+</sup> T cells was associated with the worst survival outcome (Fig. 1f). The survival association of the TLS/CD8<sup>+</sup> group was sustained in multivariate analysis adjusting for disease stage, metastasis localization, age and gender ( $P = 0.006$ , multivariate Cox regression model) (Extended Data Fig. 1b, Extended Data Table 2). Transcriptomic data showed additional differences in immunological gene signatures<sup>13</sup> (Extended Data Fig. 1c). Although TLSs were not restricted to lymph node metastases, these metastases represented the most-prevalent sample site containing TLSs (Extended Data Fig. 1d, Extended Data Table 1). To further understand the role of TLSs in tumours, we determined the spatial location of TLSs (on the tumour border or infiltrating), the number of TLSs per square millimetre and the presence of germinal-centre-like structures within

TLSs using Ki67 immunostaining (Fig. 1g, h). The location of the TLS was independent of metastatic site and, notably, tumours with infiltrative TLSs had a significantly higher frequency of melanomas with a tumour-infiltrative CD8<sup>+</sup> T cell pattern ( $P = 0.009$ , Fisher's exact test). Using survival analysis of patients with regional lymph node metastases, we found a trend for patients with tumour-infiltrative TLSs having improved survival (Extended Data Fig. 1e). In total, 44% of cases with TLS had multiple TLSs per square millimetre, and these were found only in lymph node metastases (Extended Data Fig. 1f). Moreover, we found nine cases in which canonical germinal-centre-like structures were present within TLSs (Fig. 1g). Importantly, we found cases in which TLSs containing germinal-centre-like structures coexisted with loose, non-germinal-centre-like TLSs in the same tumour (Fig. 1h). The presence in the tumour of TLSs with germinal-centre-like structures was not associated with patient outcome or the CD8<sup>+</sup> T cell infiltration pattern. In all, these data support the notion that different types of TLSs exist in individual tumours and that this is independent of the spatial location of the TLS. To reveal the molecular properties of the different T cell, B cell and tumour cell populations, we used the GeoMx digital spatial profiler (Nanostring) to perform high-plex proteomic analysis (Extended Data Table 3) with spatial resolution<sup>14</sup> (Extended Data Fig. 2a). GeoMx data from CD20<sup>+</sup> B cell populations localized in TLSs revealed two main groups, characterized by high or low expression



**Fig. 2 | B cell heterogeneity and T cell phenotypes using high-plex proteomic and scRNA-seq data.** **a**, Unsupervised hierarchical clustering of B cell populations (n = 30) from TLSs across 17 melanoma tumours. Two groups, which were independent of tumour core and patient, were clearly discerned on the basis of Ki67 expression. Proteins were filtered on the basis of scRNA-seq data<sup>18</sup>. Proteins the genes for which were expressed in B cells were included, and those genes not expressed in single B cells were excluded. **b**, T cell populations (n = 22) in or in close proximity to Ki67<sup>high</sup> or Ki67<sup>low</sup> B cell populations, respectively, were analysed for differences. Box plots of CD4 and BCL-2 show increased expression in T cells located in proximity to Ki67<sup>high</sup> B cells. P value from two-sided Wilcoxon rank-sum test. **c**, **d**, Differential analysis of T cell populations (n = 91) from 43 melanoma tumours. Proteins

were filtered on the basis of a false-discovery rate (FDR) cut-off. FDR (Benjamini–Hochberg adjustment) from P values of Kruskal–Wallis test. Box plots of selected proteins with differential expression. **e**, scRNA-seq data of CD4<sup>+</sup> and CD8<sup>+</sup> T cells in B-cell-rich and -poor tumours, respectively<sup>17</sup>. Heat map displays tumour means of 27 up- and downregulated genes, as ranked by FDR from a two-sided t-test. B-cell-poor (n = 16) and -rich (n = 16) tumours are defined as those in the lower and upper tertiles, respectively, in terms of the percentage of total cells that are B cells (<1% and >5.3%, respectively). The most significant and relevant genes are highlighted. In the box plots, the centre line is the median, the box limits are the lower and upper quartiles, and the whiskers extend to the most extreme values within 1.5 × the interquartile range (IQR).

of Ki67 (Fig. 2a, Extended Data Fig. 2b). Indeed, highly proliferating B cells may operate in germinal centres: the Ki67<sup>high</sup> tumour-associated B cells that were additionally characterized by increased CD40 expression may therefore belong to more mature TLSs<sup>15</sup>. The data provide further support for the idea that TLSs at different stages exist in the same tumour (Fig. 2a). T cells found in, or in close proximity to, TLSs with Ki67<sup>high</sup> B cells tended to have a higher proportion of CD4<sup>+</sup> cells and increased expression of BCL-2 (Fig. 2b). These T cells may therefore have undergone antigen activation that subsequently led to the upregulation of the pro-survival anti-apoptotic molecule BCL-2<sup>16</sup>. Collectively, these data support the hypothesis that these B cells and T cells belong to mature TLSs. To understand the effect of TLSs on the intratumoural T cell landscape, we analysed different properties of T cells obtained from within or in close proximity to TLSs, infiltrating T cells in tumours with TLSs and T cells from tumours without TLSs. We found increased CD4 and decreased CD8 expression in T cells from within, or in close proximity to, TLSs (Fig. 2c, d). In addition, T cells in tumours without TLSs had increased expression of TIM3, PD1 and GZMB and decreased expression of BCL-2 (Fig. 2c, d). This is consistent with a recent study that demonstrates that T cells in patients who were not responding to immune checkpoint blockade (ICB) had a dysfunctional molecular phenotype<sup>17</sup>. These findings also suggest that distinct patterns of intratumoural adaptive immune activation exist, and that these patterns

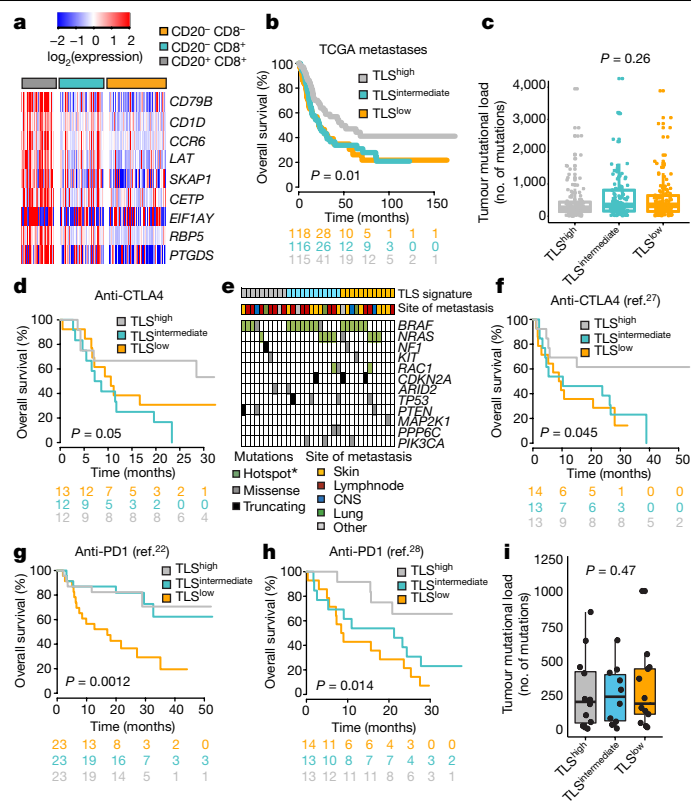
may partly be driven by TLSs. We then investigated the expression of immune markers on captured tumour cell populations. The largest difference was found when comparing tumours without an immune cell presence to other tumours. As expected, the loss of antigen presentation—via B2M and HLA-DR and decreased PDL1 expression—was found in tumours without an immune cell presence (Extended Data Fig. 2c). However, there was no difference in PDL1 expression in tumour cells between tumours with TLSs and tumours with T cells alone. We further confirmed the loss of B2M protein using immunostaining, and found that protein loss was associated with increased frequency of DNA copy number loss at the *B2M* gene locus. Moreover, we confirmed the loss of MHC using the transcriptomic data. Notably, the inflammatory state and presence of TLSs was not associated with tumour mutational burden or any specific driver-gene mutation (Extended Data Fig. 2d–g).

To gain a deeper molecular understanding of the tumour-associated B cells, we used single-cell RNA-sequencing (scRNA-seq) data. After extracting all B cells from 27 melanoma tumours used in a previous study<sup>18</sup> we then used gene sets to define activated, immature and memory B cells<sup>19</sup>, as well as plasma cells<sup>19,20</sup>. We found transcriptional evidence that a mixture of activated and immature B cells, and only a small fraction of plasma cells, are present in melanoma tumours (Extended Data Fig. 3a), which provides further support for the presence of TLSs. A fraction of single B cells expressed the class-switching and affinity



maturation gene *AICDA* or the master regulator of germinal-centre initiation, *BCL6*<sup>15</sup>. Moreover, genes important for germinal-centre initiation (*IRF4*, *POU2AF1*, *MEF2C*, *MYC*, *MEF2B*, *IRF8*, *BCL6*, *MCL1*, *TCF3*, *EBF1*, *SPIB*, *DOCK8* and *BACH2*), the germinal-centre light zone (*CD83* and *CD86*) the germinal-centre dark zone (*CXCR4*), and T cell interaction (*CD40*) were abundantly expressed in B cells<sup>15</sup> (Extended Data Fig. 3b). Thus, the transcriptional data suggest a wide range of B-cell-derived, immature-to-mature germinal-centre signals. This is consistent with the heterogeneity of TLS states observed in the immunostaining and GeoMx data. MHC class I and II molecules displayed a uniform high expression across single B cells, which suggests that B cells within TLSs are generally capable of antigen presentation. The expression of *IGLL1*, a component of the B cell receptor in pre-B cells displayed an intriguing pattern. Three clear B cell groups could be discerned; plasma cells, cells positive for *IGLL1* and *IGLL5* and cells negative for *IGLL1* and *IGLL5*. These groups could be further subdivided on the basis of *CD69* expression (Extended Data Fig. 3b). Using previously published scRNA-seq data<sup>17</sup>, we found that the fraction of *CD69*<sup>+</sup> and *IGLL5*<sup>+</sup> *CD69*<sup>+</sup> cells—and not *IGLL5*<sup>+</sup> B cells—was associated with the response to ICB (Extended Data Fig. 3c). Moreover, the *CD69*<sup>+</sup> B cell group we identified presents a more-pronounced germinal-centre-reaction phenotype than the *IGLL1*<sup>+</sup> *IGLL5*<sup>+</sup> B cell group, as *CD69* is correlated with markers of the mature germinal centre such as *CD83* and *CXCR4* (Extended Data Fig. 3d). Therefore, the observed B cell groups may reflect the maturation state of the underlying germinal-centre reaction that occurs in TLSs. By contrast, the percentage of *IGHD*<sup>+</sup> B cells ('unswitched' IgD<sup>+</sup>) and *IGHG*<sup>+</sup> B cells ('switched' IgG<sup>+</sup>) were not predictive of therapy outcome at baseline (Extended Data Fig. 3c). Collectively, these data support the presence of distinct subsets of B cells at different stages of B cell development, and their role in the response to ICB; however, further studies are needed to confirm the role of *CD69*<sup>+</sup> B cells. Finally, we investigated whether the immune microenvironment of the tumour is adapted by the presence of B cells. In single-cell data, B-cell-rich samples contained more CD4<sup>+</sup> and CD8<sup>+</sup> T cells with naive and/or memory-like characteristics (expressing *TCF7* and *IL7R*) as compared to B-cell-poor samples (Fig. 2e), suggesting an influx of naive and memory T cells to TLSs. Such memory TCF7<sup>+</sup> T cells have previously been associated with an improved response to ICB<sup>17</sup>. This is consistent with our GeoMx data, in which T cells in tumours without TLSs had an exhausted-like molecular phenotype (Fig. 2c).

Next, we used differential expression analysis to create a gene signature that reflects melanoma tumours with TLSs (Fig. 3a, Extended Data Table 4). This signature included known B-cell-specific genes such as *CD79B*. Another interesting candidate is *CCR6*, which was recently found to be upregulated in activated B cells<sup>21</sup>. Indeed, in the single-cell data from melanomas<sup>17,18</sup>, *CCR6* and *CD79B* are specifically expressed in tumour-associated B cells. The remaining genes of the signature were expressed mainly by other types of immune cell (Extended Data Fig. 3e). Similarly, the TLS-hallmark genes *CCR7*, *CXCR5* and *SELL* (which encodes CD62L) were expressed in single B cells and—to some degree—by CD4<sup>+</sup> T cells, whereas *CXCL13* is expressed predominantly by CD8<sup>+</sup> T cells (Extended Data Fig. 3f). This suggests that TLSs localized in melanoma tumours consist of B cells and other immune cells. Next, we constructed a signature from a compendium of TLS-hallmark genes (*CCL19*, *CCL21*, *CXCL13*, *CCR7*, *CXCR5*, *SELL* and *LAMP3*)<sup>11</sup>, and found that it correlates closely with our TLS signature in three datasets<sup>22–24</sup> (correlations of 0.91, 0.85 and 0.87). Further, the TLS signature correlated strongly with B cell signatures and single B cell markers. The TLS signature also correlated with signatures of T cells and other types of immune cell<sup>25,26</sup>—although not to the same extent as it did to B cell signatures (Extended Data Fig. 4a). To gain further support for the TLS signature we derived, we retrieved RNA-seq data for metastatic melanomas from The Cancer Genome Atlas (TCGA) project<sup>24</sup>. Trichotomizing the data on the basis of our TLS signature confirmed the association with patient survival (Fig. 3b, Extended Data Table 2). Analysis of matched mutation data



**Fig. 3 | TLS gene signature derived from the CD8<sup>+</sup>CD20<sup>+</sup> group predicts prognosis and response to ICB in melanoma.** **a**, Heat map of genes specifically upregulated in CD8<sup>+</sup>CD20<sup>+</sup> cases of melanoma. **b**, Kaplan–Meier analysis based on the trichotomized TLS gene signature in the melanoma metastases cohort from the TCGA ( $n = 349$  patients with available follow-up information).  $P$  value from Cox regression analysis. **c**, Mutational load across TCGA TLS groupings in **b**.  $P$  value from Kruskal–Wallis test. **d**, Kaplan–Meier analysis for overall survival in patients treated with anti-CTLA4 ( $n = 37$ ).  $P$  value from Cox regression analysis. **e**, Mutational pattern in patients treated with anti-CTLA4. **f**, Kaplan–Meier analysis for overall survival in patients treated with anti-CTLA4 ( $n = 40$ ).  $P$  value from Cox regression analysis. Data from a previous study<sup>27</sup> were used. **g**, Kaplan–Meier analysis for overall survival in patients treated with anti-PD1 ( $n = 69$ ).  $P$  value from Cox regression analysis. Data from a previous study<sup>22</sup> were used. **h**, Kaplan–Meier analysis for overall survival in patients treated with anti-PD1 ( $n = 40$ ).  $P$  value from Cox regression analysis. Data from a previous study<sup>28</sup> were used. **i**, Mutational load across the TLS grouping, using data from a previous publication<sup>28</sup>.  $P$  value from Kruskal–Wallis test. In **b**, **d**, **f**, **h**, patients were trichotomized according to high, intermediate and low expression of the TLS signature score. In the box plots, centre line is the median, the box limits are the lower and upper quartiles, and the whiskers extend to the most extreme values within  $1.5 \times \text{IQR}$ . Numbers below plots represent numbers of patients.

revealed no difference in mutational burden (Fig. 3c). Notably, samples with a TLS<sup>high</sup> signature also included non-lymph-node metastases, and—when extended to primary tumours—a small portion of the primary tumours also had a high TLS gene score (Extended Data Fig. 4b). Collectively, this confirms a prognostic role for TLS in melanoma.

Given the success of ICB in treating melanoma, we investigated the importance of tumour-associated TLSs in response to therapy (Extended Data Table 5). First, we gathered a collection of melanoma tumour biopsies from patients who were receiving CTLA4 blockade. Trichotomizing gene-expression data on the basis of the TLS signature revealed that TLS<sup>high</sup> tumours in particular were associated with significantly increased survival after CTLA4 blockade (Fig. 3d, Extended Data Fig. 5a, Extended Data Table 2). Mutation data in melanoma driver genes further supported the notion that the TLS signature is independent of tumour genetic mechanisms (Fig. 3e). We further verified the predictive

effect of the TLS signature using previously published<sup>27</sup> data from an additional cohort of 40 patients with melanoma who were receiving CTLA4 blockade (Fig. 3f). Previous studies have demonstrated tumour mutational burden as a predictive biomarker for response to ICB<sup>28</sup>. However, in this cohort of patients treated with anti-CTLA4, the TLS signature is independent of mutational load (Extended Data Fig. 5b). Moreover, the TLS signature was significantly associated with overall survival in a previously published<sup>22</sup> dataset of pretreatment samples from 69 patients who were undergoing anti-PD1 monotherapy or anti-CTLA4 and anti-PD1 combination therapy (Fig. 3g). We also observed the predictive effect of the TLS signature in a previously published<sup>28</sup> dataset of pretreatment samples from 41 patients who were treated with anti-PD1 (of whom 50% had been exposed to anti-CTLA4 before anti-PD1 treatment) (Fig. 3h, Extended Data Fig. 5c). Finally, we performed meta Cox regression analysis across the four cohorts treated with ICB, using multiple immune signatures: of these, our TLS signature performed best (Extended Data Fig. 5d). The TLS signature was also independent of tumour mutational load in the cohort treated with anti-PD1 (Fig. 3i), consistent with previous studies that have shown that immune gene signatures are not correlated with mutational load<sup>29</sup>. Although we did not observe significant differences in the TLS gene-expression score retrieved from pretreatment biopsies with regards to ‘response evaluation criteria in solid tumours’ (RECIST), we observed a notable difference in RNA-seq data from on-treatment biopsies that were collected on cycle 1 at day 29, which was confirmed in previously published cases of patients treated with anti-PD1<sup>30</sup> (Extended Data Fig. 5e, f). This indicates that TLS functionality is induced by ICB treatment in patients with a clinical response. To further determine the biological relevance of our TLS signature, we applied it to RNA-seq data from 13 additional samples of melanoma that were obtained from patients who were receiving ICB, and performed concurrent immunostaining of CD20 and CD3. The samples with the highest TLS gene score contained TLSs (as detected by CD20 immunostaining), which confirms the ability of our gene signature to predict samples with TLSs (Extended Data Fig. 5g).

In conclusion, our data provide evidence that TLSs may have a key role in sustaining an immune-responsive microenvironment. This finding opens avenues for therapeutic strategies that aim at enhancing TLS formation and function, which could result in improved clinical outcomes and responses to cancer immunotherapy.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-019-1914-8>.

1. Robert, C. et al. Pembrolizumab versus ipilimumab in advanced melanoma. *N. Engl. J. Med.* **372**, 2521–2532 (2015).
2. Gopalakrishnan, V. et al. Gut microbiome modulates response to anti-PD-1 immunotherapy in melanoma patients. *Science* **359**, 97–103 (2018).

3. Cristescu, R. et al. Pan-tumor genomic biomarkers for PD-1 checkpoint blockade-based immunotherapy. *Science* **362**, eaar3593 (2018).
4. Ladányi, A. et al. Prognostic impact of B-cell density in cutaneous melanoma. *Cancer Immunol. Immunother.* **60**, 1729–1738 (2011).
5. Messina, J. L. et al. 12-Chemokine gene signature identifies lymph node-like structures in melanoma: potential for patient selection for immunotherapy? *Sci. Rep.* **2**, 765 (2012).
6. Cipponi, A. et al. Neogenesis of lymphoid structures and antibody responses occur in human melanoma metastases. *Cancer Res.* **72**, 3997–4007 (2012).
7. Sautès-Fridman, C., Petitprez, F., Calderaro, J. & Fridman, W. H. Tertiary lymphoid structures in the era of cancer immunotherapy. *Nat. Rev. Cancer* **19**, 307–325 (2019).
8. Petitprez, F. A. d. R. et al. B cells are associated with survival and immunotherapy response in sarcoma. *Nature* <https://doi.org/10.1038/s41586-019-1906-8> (2020).
9. Helmink, B. A. et al. B cells and tertiary lymphoid structures promote immunotherapy response. *Nature* <https://doi.org/10.1038/s41586-019-1922-8> (2020).
10. Mihm, M. C., Jr & Mulé, J. J. Reflections on the histopathology of tumor-infiltrating lymphocytes in melanoma and the host immune response. *Cancer Immunol. Res.* **3**, 827–835 (2015).
11. Dieu-Nosjean, M. C., Goc, J., Giraldo, N. A., Sautès-Fridman, C. & Fridman, W. H. Tertiary lymphoid structures in cancer and beyond. *Trends Immunol.* **35**, 571–580 (2014).
12. Germain, C., Gnjatich, S. & Dieu-Nosjean, M. C. Tertiary lymphoid structure-associated B cells are key players in anti-tumor immunity. *Front. Immunol.* **6**, 67 (2015).
13. Bindea, G. et al. Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. *Immunity* **39**, 782–795 (2013).
14. Amaria, R. N. et al. Neoadjuvant immune checkpoint blockade in high-risk resectable melanoma. *Nat. Med.* **24**, 1649–1654 (2018).
15. De Silva, N. S. & Klein, U. Dynamics of B cells in germinal centres. *Nat. Rev. Immunol.* **15**, 137–148 (2015).
16. Rogers, P. R., Song, J., Gramaglia, I., Killeen, N. & Croft, M. OX40 promotes Bcl-xL and Bcl-2 expression and is essential for long-term survival of CD4 T cells. *Immunity* **15**, 445–455 (2001).
17. Sade-Feldman, M. et al. Defining T cell states associated with response to checkpoint immunotherapy in melanoma. *Cell* **175**, 998–1013 (2018).
18. Jerby-Arnon, L. et al. A cancer cell program promotes T cell exclusion and resistance to checkpoint blockade. *Cell* **175**, 984–997 (2018).
19. Angelova, M. et al. Characterization of the immunophenotypes and antigenomes of colorectal cancers reveals distinct tumor escape mechanisms and novel targets for immunotherapy. *Genome Biol.* **16**, 64 (2015).
20. Tarte, K., Zhan, F., De Vos, J., Klein, B. & Shaughnessy, J. Jr. Gene expression profiling of plasma cells and plasmablasts: toward a better understanding of the late stages of B-cell differentiation. *Blood* **102**, 592–600 (2003).
21. Suan, D. et al. CCR6 defines memory B cell precursors in mouse and human germinal centers, revealing light-zone location and predominant low antigen affinity. *Immunity* **47**, 1142–1153 (2017).
22. Gide, T. N. et al. distinct immune cell populations define response to anti-PD-1 monotherapy and anti-PD-1/anti-CTLA-4 combined therapy. *Cancer Cell* **35**, 238–255 (2019).
23. Cirenajwis, H. et al. NF1-mutated melanoma tumors harbor distinct clinical and biological characteristics. *Mol. Oncol.* **11**, 438–451 (2017).
24. Cancer Genome Atlas Network. Genomic classification of cutaneous melanoma. *Cell* **161**, 1681–1696 (2015).
25. Becht, E. et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol.* **17**, 218 (2016).
26. Tirosh, I. et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* **352**, 189–196 (2016).
27. van Allen, E. M. et al. Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. *Science* **350**, 207–211 (2015).
28. Riaz, N. et al. Tumor and microenvironment evolution during immunotherapy with nivolumab. *Cell* **171**, 934–949 (2017).
29. Lauss, M. et al. Mutational and putative neoantigen load predict clinical benefit of adoptive T cell therapy in melanoma. *Nat. Commun.* **8**, 1738 (2017).
30. Roh, W. et al. Integrated molecular analysis of tumor biopsies on sequential CTLA-4 and PD-1 blockade reveals markers of response and resistance. *Sci. Transl. Med.* **9**, eaah3560 (2017).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2020

## Methods

No statistical methods were used to predetermine sample size. The experiments were not randomized and investigators were not blinded to allocation during experiments and outcome assessment.

### Patient material

This study was approved by the Regional Ethics Committee at Lund University (Dnr. 191/2007 and 101/2013). The sample cohort, representing a population-based retrospective collection ( $n = 177$ ), was obtained at the Department of Surgery at Skåne University Hospital.

Overall, 104 patients had regional metastatic disease, 50 distant disease and 19 local disease. Four patients were of unknown stage. This is a historical cohort, collected between 2000 and 2012. As such, the cohort is suitable for prognostic studies. A summary of the patient characteristics is provided in Extended Data Table 1.

We also collected paraffin-embedded tumour tissue from 119 patients; 37 of these patients had received anti-CTLA4 as first-line therapy. Tumour tissue was collected from these patients in Denmark, and available biopsies were obtained a maximum of six months before the start of therapy. This study was approved by the regional ethical committee (H-15010200). DNA and RNA were extracted using the Qia-gen FFPE AllPrep procedure, as previously described<sup>31</sup>.

We retrieved frozen tumour tissue from 13 patients who were undergoing anti-PD1 therapy at Skåne University Hospital, under ethical approval Dnr. 101/2013. RNA-seq and data analysis was performed as previously described<sup>29</sup>.

### High-plex proteomic analysis

We used the Nanostring GeoMx platform for high-plex proteomic analysis with spatial resolution, as previously described<sup>14</sup>. Two 5- $\mu$ m tissue microarray slides were used. Antibodies against CD3, CD20, DAPI, and PMEL and S100B were used for immunofluorescence, which was subsequently used for region of interest selection and UV masking. Digital counts from barcodes corresponding to protein probes (in total 60 immune-related proteins) were analysed as follows: raw counts were first normalized with internal spike-in controls (ERCC) to account for system variation. To control for nonspecific antibody binding, values were further normalized by a linear scaling factor to obtain IgG control counts of 1 for each region of interest. To reduce background noise, values below 3 were set to 1 and the data were log<sub>2</sub>-transformed. Data are provided in Supplementary Information.

### Immunohistochemistry

Tissue microarrays were constructed using, on average, three 1-mm cores per tumour in an attempt to obtain a representative picture of the tumour. The tissue block was cut in 4- $\mu$ m sections, and then dried at 60 °C for 1 h. The paraffin-embedded sections were deparaffinized and pre-treated in the PT-Link (DAKO) with target retrieval solution buffer pH 9. The following steps (except for the primary antibody staining) were performed in the DAKO staining equipment (Autostainer plus) with DAKO kit K8010 solutions: peroxidase block (5 min), EnVision HRP-conjugated polymers (30 min), DAB substrate–chromogen solution (2 × 5 min) and counterstaining with haematoxylin (4 min). Between each step, the sections were rinsed with washing buffer. Finally, the sections were dehydrated and mounted with PERTEX mounting medium (ref. 00811) (Histolab). The primary antibodies used were all from Agilent/DAKO: CD3 (A0452) in 1:200 dilution, CD8 (M7103) in 1:100 dilution, MITF (Clone C5), B2M (A0072), Ki67 (MIB-1) in 1:500 dilution and CD20 (M0755) in 1:400 dilution. SOX10 was performed in the clinical routine laboratory of clinical pathology (Skåne University Hospital) using the mouse monoclonal IgG1 (clone BC34, Biocare Medical) antibody.

### Immunofluorescence staining

Initially, the cells from snap-frozen tumours known to have TLSs were incubated in ice-cold acetone for 10 min and washed in PBS. All the

following steps were performed in a humidified chamber. Unspecific binding sites were masked with PBS + 3% BSA for 90 min at room temperature. Mouse-anti-CD20 (1:200, 00064779, DAKO), rabbit-anti-CXCR5 (1:200, 3180237-9, Abcam) and rabbit-anti-CXCL13 (1:200, NBP2-1604155, Novus Biologicals) were applied overnight at 4 °C. Donkey-anti-mouse-AF488 and goat-anti-rabbit-AF546 was applied 1:1,000 in PBS + 1% BSA for 90 min at room temperature, followed by mounting with DAPI-containing mounting medium (Vector Laboratories). Fluorescence images were acquired with an Olympus BX63 microscope, DP80 camera and cellSens Dimension v.1.12 software (Olympus).

### Bioinformatic and statistical analyses

**Datasets.** Microarray expression data were generated using the Illumina HT12 arrays, and have been used in a previous publication<sup>23</sup>; they are deposited in Gene Expression Omnibus, accession number GSE65904. Mutation data were generated using a sequencing panel targeting 1,550 cancer genes, as previously described<sup>23</sup>, and copy number data were derived from the corresponding raw sequencing data using Contra version 2.0.3<sup>32</sup> with segmentation using GLAD<sup>33</sup>.

RNA-seq data of metastatic melanomas from TCGA (level 3, release 3.1.14.0) were downloaded from the data portal, quantile-normalized and log-transformed as log<sub>2</sub>(data + 1).

The PD1-treatment RNA-seq data from ref. <sup>22</sup> were downloaded as fastq files from the European Nucleotide Archive (PRJEB23709) and fragments per kilobase of transcript per million mapped reads (FPKM) values were retrieved using HISAT and Stringtie<sup>34</sup>. The data were reduced to protein-coding genes, samples were quantile-normalized and the data were log-transformed as log<sub>2</sub>(data + 1). Previously published PD1 inhibitor-treatment RNA-seq data<sup>28</sup> were downloaded as count data ('CountData.BMS038.txt') with annotations from [https://github.com/riazn/bms038\\_analysis/tree/master/data](https://github.com/riazn/bms038_analysis/tree/master/data). The data were reduced to protein-coding genes and normalized for transcript lengths using exon annotations from the R package TxDb.Hsapiens.UCSC.hg19.knownGene, subsequently transformed to transcripts per million (TPM) and quantile-normalized. The data were log-transformed as log<sub>2</sub>(data + 2) – 1. Previously published NanoString gene-expression data<sup>30</sup> were downloaded from the respective supplementary table. Previously published CTLA4 inhibitor-treatment data<sup>27</sup> were received from the authors as reads per kilobase of transcript per million mapped reads (RPKM) values; the data were quantile-normalized and log-transformed as log<sub>2</sub>(data + 1). scRNA-seq data were retrieved from Gene Expression Omnibus accessions GSE115978 and GSE120575, protein-coding genes were kept and cells with less than 1,700 or 1,000 genes expressed >0 were removed, respectively. Data for B cells were extracted, and quantile-normalized. For GSE115978, the available B cell definition was used; for GSE120575, no B cell definition was available and B cells were defined as CD19 > 2.

We generated gene-expression profiles from 119 formalin-fixed paraffin-embedded (FFPE) samples using Affymetrix Clariom D microarrays. The hybridized FFPE material constituted three separate retrospective studies, including the 37 pre-ipilimumab treatment samples analysed in this study. Principal component (PC) analysis informed us that this FFPE-derived data was greatly affected by sample degradation. We therefore reduced the data to probesets mapping to the 3' untranslated region (UTR) of curated RefSeq transcripts; using PRINCIPAL categories from APPRIS<sup>35</sup>, we obtained 33,111 probesets in the 3' UTRs of the principal gene isoforms. We further selected the two cohorts from the same Danish site, and removed one sample with a sample median expression < 0 and 7 samples with a median control exon ('HTA2-pos' probes) minus median control intron ('HTA2-neg' probes) expression < 1. The remaining probesets were filtered for being expressed, by keeping probesets that were above the median control intron expression in at least 90% of samples (19,990 probesets). The most-varying probeset for each protein-coding gene was kept (10,197 genes), and quantile

normalization was applied. As the data was still affected by degradation, PC1 and PC2 of the data were removed using R package *swamp*<sup>36</sup>, an offset of 1.5 was added to revert negative values, and the 5,000 genes with the largest variation were kept. Gene-expression data of the 37 samples of cutaneous melanoma with ipilimumab pretreatment were extracted for this study (Supplementary Information). Additional data and codes are available from the corresponding author upon request.

**TLS signature.** To derive the TLS-signature genes, we performed SAM analysis<sup>37</sup> to identify genes overexpressed in CD8<sup>+</sup>CD20<sup>+</sup> versus CD8<sup>+</sup> groups and subtracted the genes overexpressed in CD8<sup>+</sup> versus double-negative groups (Extended Data Table 4). For each dataset, the signature genes that were present were extracted. Failed genes were defined as having an average Pearson correlation <0.15 to the other signature genes, and were excluded. The signature score was calculated as the mean gene expression. For survival analyses, the signature score was divided into equally sized tertiles.

**Statistical analyses.** Fisher's exact test was used for comparison of categorical variables. Pearson correlation was used for comparison of numerical variables. The *t*-test or Wilcoxon test and analysis of variance (ANOVA) were used for group comparisons of two or more than two groups, respectively. Owing to outliers, we used the Kruskal–Wallis test for the association of mutational load with the immunohistochemical groups. For univariate and multivariate survival analyses, we used Cox regression from the survival package. All bioinformatical analyses were done in R. All tests were two-sided. All box plots are depicted with the centre line representing the median, the box limits representing the lower and upper quartiles, and the whiskers extending to the most extreme values within 1.5× IQR.

## Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

## Data availability

All relevant data are available and are included as Source Data. Digital spatial-profiling data used in Fig. 2 and gene-expression microarray data from Danish patients treated with anti-CTLA4 are available as Source Data. Data from public repositories were accessed from GSE65904

(ref.<sup>23</sup>), TCGA data portal SKCM level 3 release 3.1.14.0, PRJEB23709 (ref.<sup>22</sup>), [https://github.com/riazn/bms038\\_analysis/tree/master/data](https://github.com/riazn/bms038_analysis/tree/master/data), GSE115978 (ref.<sup>18</sup>) and GSE120575 (ref.<sup>17</sup>). Any other relevant data and code can be obtained from the corresponding authors upon reasonable request.

31. Harbst, K. et al. Molecular profiling reveals low- and high-grade forms of primary melanoma. *Clin. Cancer Res.* **18**, 4026–4036 (2012).
32. Li, J. et al. CONTRA: copy number analysis for targeted resequencing. *Bioinformatics* **28**, 1307–1313 (2012).
33. Hupé, P., Stransky, N., Thiery, J. P., Radvanyi, F. & Barillot, E. Analysis of array CGH data: from signal ratio to gain and loss of DNA regions. *Bioinformatics* **20**, 3413–3422 (2004).
34. Pertea, M., Kim, D., Pertea, G. M., Leek, J. T. & Salzberg, S. L. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat. Protocols* **11**, 1650–1667 (2016).
35. Rodriguez, J. M. et al. APPRIS: annotation of principal and alternative splice isoforms. *Nucleic Acids Res.* **41**, D110–D117 (2013).
36. Lauss, M. et al. Monitoring of technical variation in quantitative high-throughput datasets. *Cancer Inform.* **12**, 193–201 (2013).
37. Tusher, V. G., Tibshirani, R. & Chu, G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl Acad. Sci. USA* **98**, 5116–5121 (2001).

**Acknowledgements** The study was supported by the Swedish Cancer Society, the Swedish Research Council, BioCARE, the Berta Kamprad Foundation, the King Gustaf V Jubilee foundation, Mats Paulsson's foundation, Stefan Paulsson's foundation and governmental funding for healthcare research (ALF). The Danish Cancer Society, the Aase and Einar Danielsen's Fund and the Capital Region of Denmark Research Foundation. The research leading to these results has received funding from the European Community's Horizon 2020 Framework Programme for Research and Innovation (H2020-MSCA-ITN-2014) under Grant Agreement no. 247634.

**Author contributions** G.J. conceived and supervised the study. R.C., M.L. and G.J. analysed and drafted text. R.C., B.P., K.L. and K.J. generated immunostaining data. R.C., A.S., I.J., B.P. and G.J. analysed immunostaining data. R.C., K.P. and G.J. generated and analysed immunofluorescence data. A.v.S. and S.W. generated digital spatial-profiling data. M.L. and G.J. analysed digital spatial-profiling data. R.C., M.L., S.M., K.H. and G.J. performed statistical analyses. M.L. and G.J. performed bioinformatical analyses. M.L. analysed scRNA-seq data. J.V.-C. generated RNA-seq data. M.L., A.S., M.D., M.S.L., I.J., B.P., K.H., J.V.-C., A.v.S., K.L., S.W., K.J., K.P., D.S., J.A.W. and G.J. interpreted data. M.D., M.S.L., H.O., C.I., K.I., H.S., L.B., A.C. and I.M.S. collected clinical specimens and clinical data. All authors approved and read the final draft.

**Competing interests** S.W. and A.v.S. are employees of Nanostring Inc. and declare that there are competing interests. All other authors declare no conflict of interest.

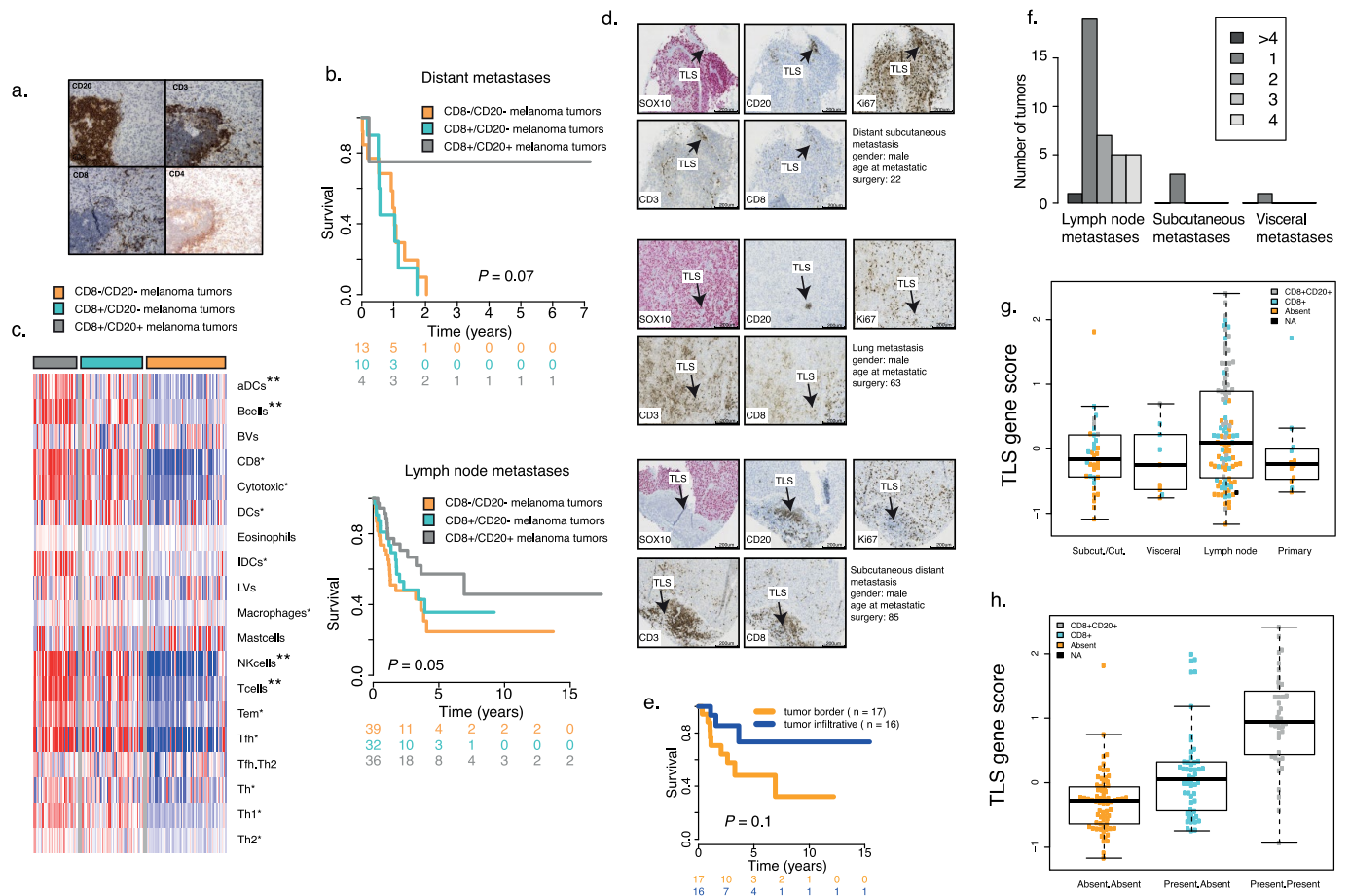
## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41586-019-1914-8>.

**Correspondence and requests for materials** should be addressed to G.J.

**Peer review information** Nature thanks James J. Mulé, Caroline Robert and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

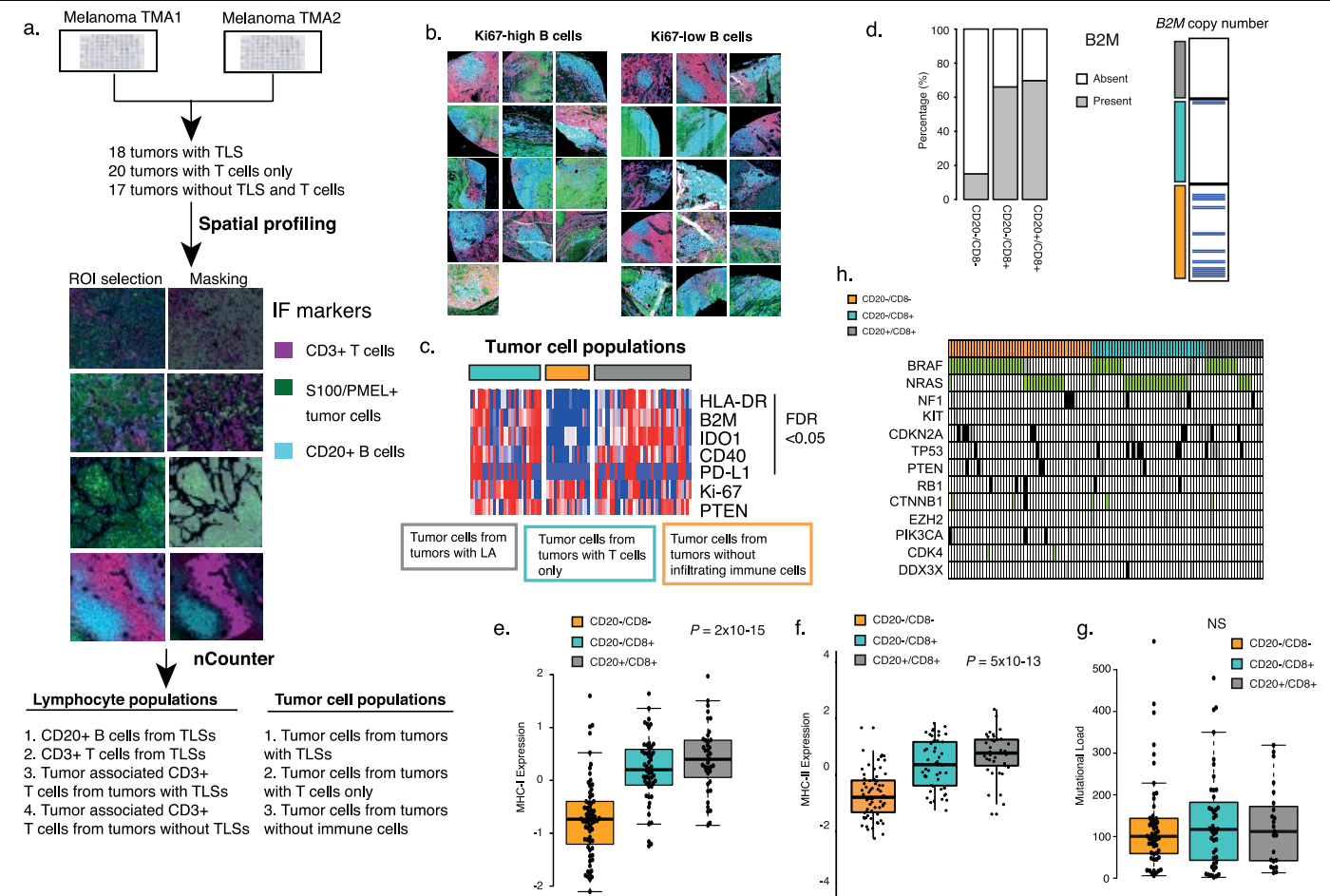
**Reprints and permissions information** is available at <http://www.nature.com/reprints>.



**Extended Data Fig. 1 | Characterization of TLSs in melanoma tumours.** **a**, CD20 (B cells), CD3 (T cells), CD8 (CD8<sup>+</sup> T cells) and CD4 (CD4<sup>+</sup> T cells) immunostaining in a representative melanoma with a TLS ( $n = 44$  cases with TLS in the cohort of 177 cases). **b**, Subset survival analysis using CD8 and CD20 immunostaining in distant and lymph node metastases separately.  $n = 27$  and 97 patients with available follow-up information, respectively.  $P$  values from Cox regression. **c**, Gene-expression characterization of the three groups using previously described signatures<sup>13</sup>. aDCs, activated dendritic cells; BVs, blood vessels; DCs, dendritic cells; IDCs, immature dendritic cells; LVs, lymph vessels; Macrophages, Mast cells; NK cells; Tem, T effector memory cells; Tfh, T follicular helper cells; Tfh, Th2, T follicular helper 2 cells; Th, T helper cell; Th1, T helper 1 cell; Th2, T helper 2 cell.

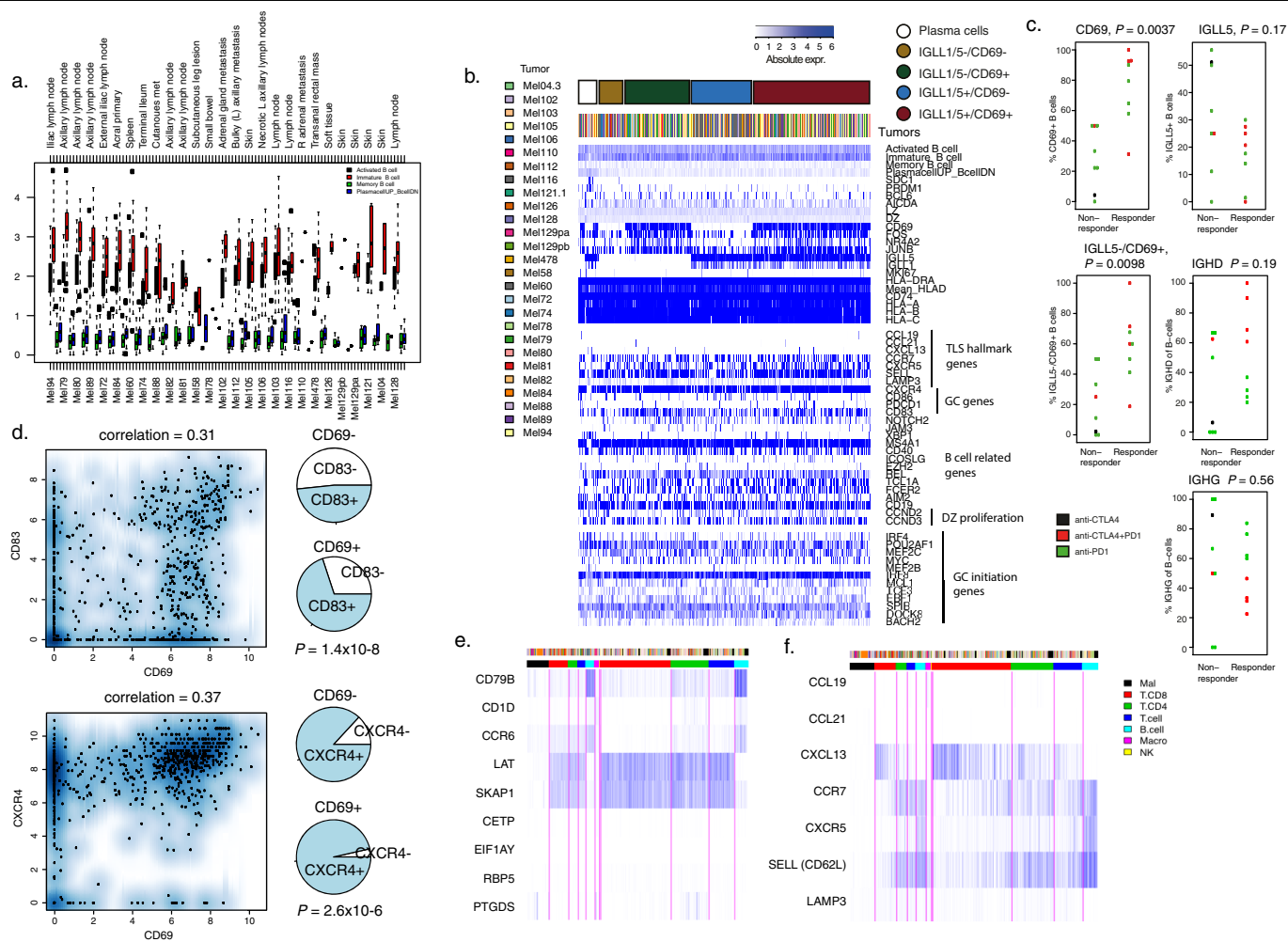
2 cell. **d**, CD20, CD3, CD8, Ki67 and SOX10 immunostainings in three distant metastases. Arrows indicate the TLS. **e**, Survival analysis of 33 patients with TLS-containing tumours from regional lymph node metastases, stratified according to whether the TLS is located at the tumour border or is tumour-infiltrative.  $P$  value from Cox regression. **f**, Bar plot showing quantification of TLSs in tumours. Numbers in the box corresponds to TLSs per square millimetre. **g**, TLS gene score and type of lesion.  $n = 159$  tumours. **h**, TLS score and immunological group.  $n = 159$  tumours. In the box plots, the centre line represents the median, the box limits represent the lower and upper quartiles, and the whiskers extend to the most extreme values within  $1.5 \times \text{IQR}$ . Numbers below the graphs represent numbers of patients.





**Extended Data Fig. 2 | High-plex proteomic analysis using the GeoMx assay and genomic characterization of tumours containing TLSs.** **a**, Workflow of the GeoMx assay. **b**, Immunofluorescence imaging of TLSs in tumour samples used in the GeoMx analysis. TLSs are sorted according to the unsupervised clustering of the high-plex proteomic data, performed on the different B cell populations. Pink, CD3<sup>+</sup> T cells; green, tumour cells positive for PMEL and/or S100B; cyan, CD20<sup>+</sup> B cells. For Ki67<sup>high</sup> 13 of 13 TLSs are displayed, and for Ki67<sup>low</sup> 15 of 17 TLSs are displayed. **c**, GeoMx data from 83 captured tumour cell regions. FDRs are from Kruskal–Wallis test, adjusted for multiple testing using the Benjamini–Hochberg method. **d**, Left, B2M immunostaining shows a

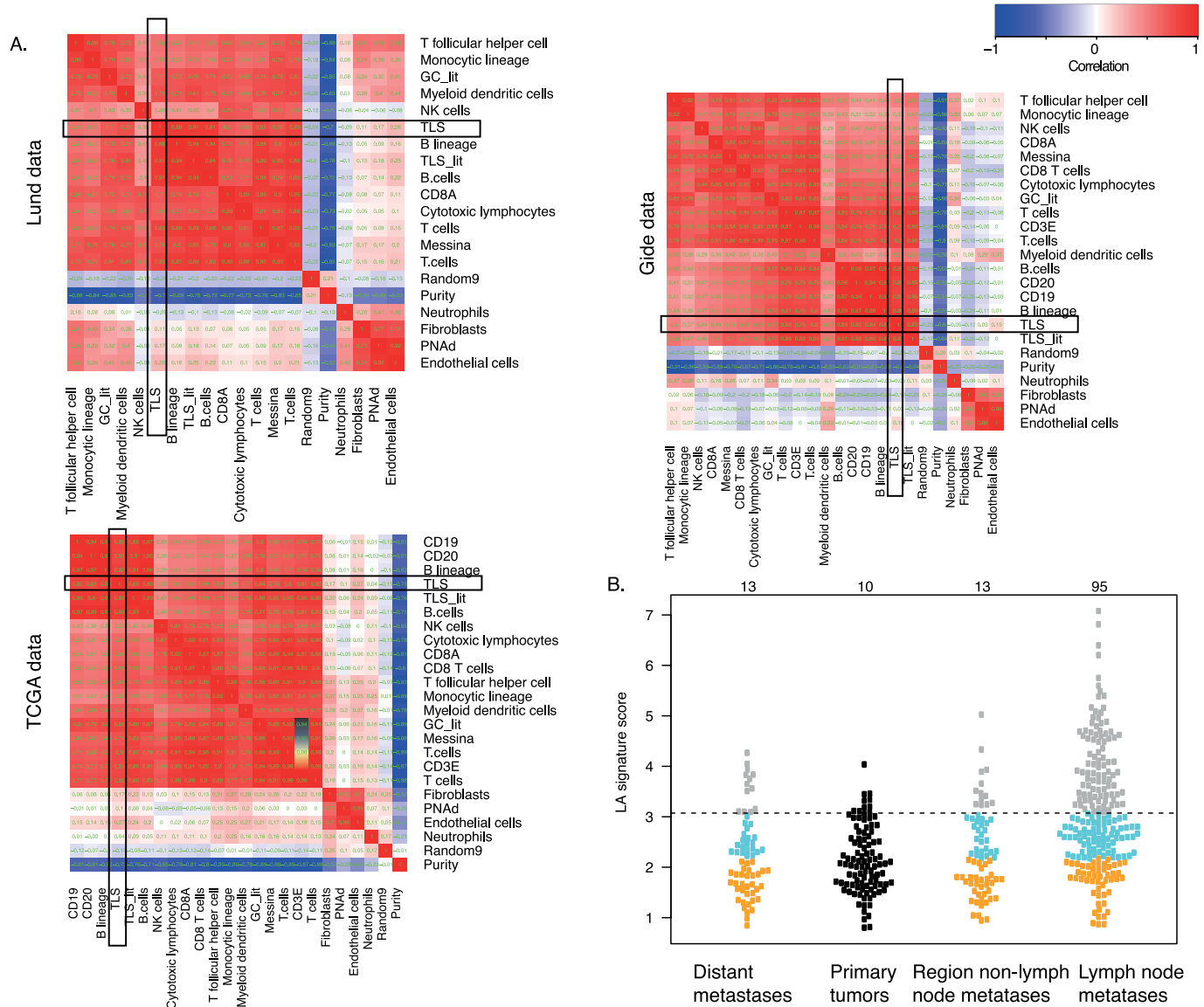
significant difference between CD8/CD20 groups.  $P = 1 \times 10^{-11}$ , Fisher's exact test,  $n = 172$  tumours). Right, plot shows B2M copy number status (blue = loss).  $P = 0.002$ , FDR adjustment for multiple comparisons = 0.007, Fisher's exact test,  $n = 127$  tumours. **e–g**, MHC-I (**e**) and MHC-II (**f**) expression ( $n = 160$  tumours,  $P$  value from ANOVA) and mutational load (**g**) ( $n = 118$  tumours, Kruskal–Wallis test) in relation to immunological groupings. **h**, Mutation heat map of melanoma-relevant genes in relation to immunological grouping. In the box plots, the centre line represents the median, the box limits represent the lower and upper quartiles, and the whiskers extend to the most extreme values within 1.5× IQR.



### Extended Data Fig. 3 | scRNA-seq analysis of tumour-associated B cells.

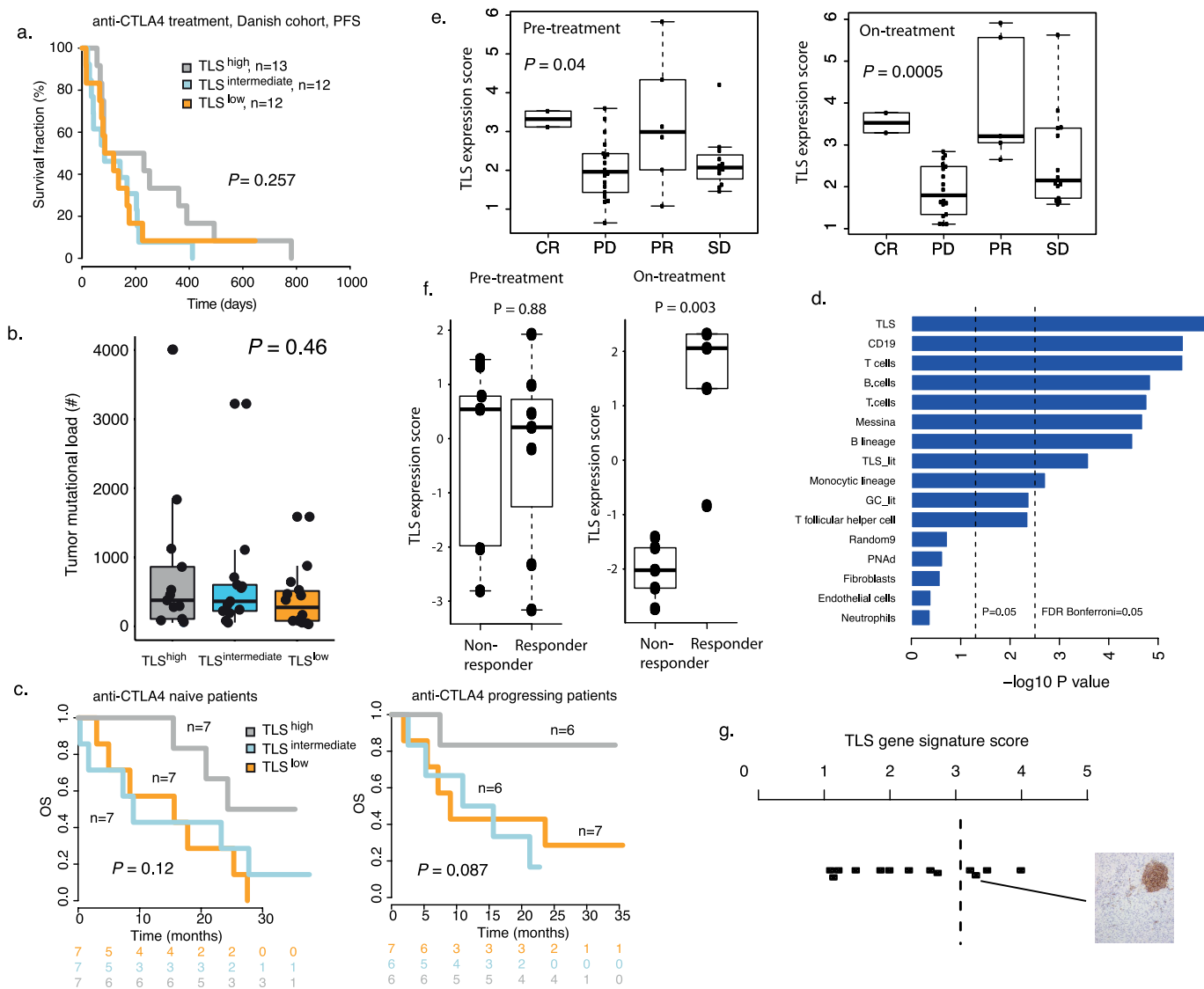
**a**, Box plots of gene-expression scores, based on different B cell developmental states in 812 B cells from 27 tumours from a previous study<sup>18</sup>. **b**, Heat map of selected genes across all 812 B cells. *IGLL5* and *CD69* were two of the five genes with highest expression variation across all B cells. The heat map is sorted on *IGLL5* and *CD69* expression, excluding the cells that displayed increased expression of the plasma-cell signature. Genes showing a Pearson correlation  $>0.4$  to *IGLL5* or *CD69* expression are also indicated. *SDCI* and *PRDM1* mark plasma cells, *BCL6* and *AICDA* mark germinal centres, *HLA-DRA* mark MHC-II and *HLA-A*, *HLA-B* and *HLA-C* mark MHC-I. TLS-hallmark genes, germinal-centre-related genes and other B cell genes are also indicated. **c**, Extracting the single B cell RNA-seq data from a previous study<sup>17</sup> using pretreatment samples ( $n=16$ ). The fraction of *CD69*<sup>+</sup> B cells was higher in responders to ICB than in nonresponders ( $n=8$ ), but the fraction of *IGLL5*<sup>+</sup> B cells was not. The fraction of *IGLL5*<sup>+</sup> *CD69*<sup>+</sup> cells was also higher in responders. Plots of fraction of *IGHG*<sup>+</sup> B cells in relation to response to ICB therapy. *P* values from two-sided Wilcoxon test. In **a**, the centre lines in the box plot represent the median, the box limits represent the lower and upper quartiles, and the whiskers extend to the most extreme values within  $1.5 \times$  IQR. **d**, Pearson correlation between expression of *CD69* and germinal centre genes (*CD83* and *CXCR4*) in data from a previous study<sup>18</sup>. Pie charts display the fact that the fraction of *CD83*<sup>+</sup> and *CXCR4*<sup>+</sup> B cells is increased among *CD69*<sup>+</sup> B cells. Expression  $>1$  was used as a cut-off for being present. Seven hundred and fifty-three B cells without a present plasma-cell signature were analysed. *P* value from two-sided Fisher's exact test. **e, f**, Heat map of gene-expression values corresponding to our TLS signature (**e**) and TLS-hallmark genes from the literature (**f**). Blue corresponds to increased expression. Mal., malignant cells. In **e, f**, single cells from the seven cell types on the left are from ref.<sup>18</sup>, and from the four cell types on the right are from ref.<sup>17</sup>.

*IGHG*<sup>+</sup> B cells in relation to response to ICB therapy. *P* values from two-sided Wilcoxon test. In **a**, the centre lines in the box plot represent the median, the box limits represent the lower and upper quartiles, and the whiskers extend to the most extreme values within  $1.5 \times$  IQR. **d**, Pearson correlation between expression of *CD69* and germinal centre genes (*CD83* and *CXCR4*) in data from a previous study<sup>18</sup>. Pie charts display the fact that the fraction of *CD83*<sup>+</sup> and *CXCR4*<sup>+</sup> B cells is increased among *CD69*<sup>+</sup> B cells. Expression  $>1$  was used as a cut-off for being present. Seven hundred and fifty-three B cells without a present plasma-cell signature were analysed. *P* value from two-sided Fisher's exact test. **e, f**, Heat map of gene-expression values corresponding to our TLS signature (**e**) and TLS-hallmark genes from the literature (**f**). Blue corresponds to increased expression. Mal., malignant cells. In **e, f**, single cells from the seven cell types on the left are from ref.<sup>18</sup>, and from the four cell types on the right are from ref.<sup>17</sup>.



**Extended Data Fig. 4 | Comparison of the derived TLS gene signature to other immune signatures. a,** Pearson correlation plots of the data from the cohort obtained at Skåne University Hospital, Lund (top,  $n = 160$ ), data from cases of melanoma metastasis in the TCGA (bottom,  $n = 363$ ) and baseline data from a previous publication<sup>22</sup> (right,  $n = 69$ ). Black box indicates the TLS

signature. All signatures are taken from refs.<sup>11,19,26</sup>. Red, positive correlation; blue, negative correlation. **b,** TLS gene-signature scores in primary tumours in comparison to distant and lymph node metastases. The number of tumours assigned to the TLS<sup>high</sup> category is indicated above the plot.



#### Extended Data Fig. 5 | TLS gene signature in cohorts treated by ICB.

**a.** Progression-free survival (PFS) and TLS gene signature in the Danish cohort of patients treated with anti-CTLA4.  $P$  value from Cox regression. **b.** TLS gene signature in relation to tumour mutational load, in data from a previous publication<sup>27</sup> (n = 40 melanoma tumours).  $P$  value from Kruskal–Wallis test. **c.** Survival analyses on data from a previous study<sup>28</sup>, stratified according to whether patients are naive to anti-CTLA4 treatment or have progressed on anti-CTLA4.  $P$  values from Cox regression. **d.** Meta Cox regression analysis across the four cohorts treated using ICB (n = 186).  $P$  values from Cox regression adjusted for study. **e.** TLS gene signature of pretreatment (n = 16)

and on-treatment samples (n = 10) in relation to therapy response in data from a previous publication<sup>30</sup>.  $P$  value from two-sided  $t$ -test. **f.** TLS gene signature of pretreatment (n = 38) and on-treatment (n = 39) samples in relation to RECIST response in data from a previous study<sup>28</sup>.  $P$  value from ANOVA test. **g.** TLS gene signature score in 13 melanoma tumours that were also stained for CD20 protein. As an example, the tumour with the third highest score had TLSs. The two top tumours also had TLSs, whereas the other tumours did not. In the box plots in **b**, **d**, **e**, centre lines represent the median, the box limits represent the lower and upper quartiles, and the whiskers extend to the most extreme values within 1.5 × IQR.

**Extended Data Table 1 | Clinical features of the 177-patient cohort, shown in correlation with CD8/CD20 immunological grouping**

	ENTIRE COHORT (N=177)*	CD8 <sup>+</sup> /CD20 <sup>+</sup> (N=44)	CD8 <sup>+</sup> /CD20 <sup>-</sup> (N=57)	CD8 <sup>-</sup> /CD20 <sup>-</sup> (N=74)	P VALUE
<b>PATIENT CHARACTERISTICS</b>					
<b>GENDER N (%)</b>					0.08
<b>MALE</b>	101 (57)	30 (70)	27 (47)	43 (60)	
<b>FEMALE</b>	72 (41)	13 (30)	30 (53)	29 (40)	
<b>NA</b>	4 (2)				
<b>AGE AT DIAGNOSIS MEDIAN (RANGE)</b>	65 (22-91)	66 (22-85)	64.5 (30-88)	65 (25-91)	
<b>TUMOR CHARACTERISTICS</b>					
<b>STAGE</b>					0.003
<b>II</b>	19	-	8	11	
<b>III</b>	104	35	29	39	
<b>IV</b>	50	7	20	23	
<b>NA</b>	4	2	-	1	
<b>METASTASIS TYPE</b>					0.003
<b>LYMPH NODE</b>	113	38	33	41	
<b>SUBCUTANEOUS</b>	35	3	12	20	
<b>VISCERAL</b>	10	1	5	4	
<b>PRIMARY TUMOR</b>	15	-	7	8	
<b>NA</b>	4	2	-	1	
<b>CD8 IHC</b>					
<b>INFILTRATIVE N (%)</b>	58 (33)	28 (64)	30 (53)	-	
<b>CLUSTERED N (%)</b>	43 (24)	16 (36)	27 (47)	-	
<b>ABSENT N (%)</b>	74 (42)		-	74 (100)	
<b>NA N (%)</b>	2 (1)				
<b>CD3 IHC</b>					
<b>INFILTRATIVE N (%)</b>	59 (33)	27 (61)	30 (53)	2 (3)	
<b>CLUSTERED N (%)</b>	52 (29)	17 (39)	27 (47)	8 (11)	
<b>ABSENT N (%)</b>	64 (36)	-	-	64 (86)	
<b>NA N (%)</b>	2 (1)				
<b>CD20 IHC</b>					
<b>PRESENT N (%)</b>	44 (25)	44 (100)	0 (0)	0 (0)	
<b>ABSENT N (%)</b>	131 (74)	0 (0)	57 (44)	74 (56)	
<b>NA N (%)</b>	2 (1)				
<b>PRIMARY TUMOR CHARACTERISTIC</b>					
<b>HISTOLOGICAL SUBTYPE</b>					0.34
<b>UNKNOWN PRIMARY N (%)</b>	26 (15)	7 (16)	7 (12)	12 (16)	
<b>SSM</b>	36 (20)	10 (23)	13 (23)	13 (18)	
<b>NM</b>	57 (33)	15 (34)	21 (37)	21 (28)	
<b>OTHER</b>	17 (10)	2 (4)	3 (5)	12 (16)	
<b>NA</b>	39 (22)	10 (23)	13 (23)	16 (22)	

P values from Fisher's exact test, not adjusted for multiple testing.

\*CD8 and CD20 status was missing for two patients.



Extended Data Table 2 | Univariate and multivariate Cox regression model analysis of immunological groupings in melanoma

Univariate analysis			Multivariate analysis*		
	HR (CI)	P-value	HR (CI)	P-value	
CD8 infiltrative clustered absent	1	-			
	1.21 (0.65-2.26)	0.54			
	2.03 (1.20-3.43)	<b>0.009</b>			
CD20 present absent	1	-			
	2.18 (1.24-3.82)	<b>0.007</b>			
CD8/CD20					
CD8+CD20+	1	-	1	-	
CD8+CD20-	1.76 (0.93-3.35)	0.08	1.75 (0.91-3.37)	0.09	
CD8-CD20-	2.54 (1.40-4.61)	<b>0.002</b>	2.60 (1.42-4.77)	<b>0.002</b>	
Stage*					
II	1	-	1	-	
III	1.89 (0.58-6.07)	0.28	2.38 (0.73-7.73)	0.14	
IV	7.84 (2.38-25.73)	<b>0.0007</b>	9.39 (2.84-31.04)	<b>0.0002</b>	

Multivariate IHC marker and Stage						
	CD8	CD20	CD8/CD20	Stage	Metastasis type	Age
Model 1	0.009			4x10 <sup>-9</sup>		
Model2		0.004		1x10 <sup>-8</sup>		
Model3			0.005	9x10 <sup>-9</sup>		
Model4			0.006		0.02	
Model5			0.005		0.03	0.70
Model6*			<b>0.006</b>	8x10 <sup>-7</sup>	0.68	0.69

TCGA multivariate cox analysis				
	TLS signature	Source site**	Age	Gender
Metastases only	<b>0.03</b>	0.02	0.01	0.56
All tumors including primary tumors	<b>0.01</b>	0.05	0.02	0.53

Hazard Ratios in TCGA And ICB Cohorts		
	HR (95% CI)	p-value
TCGA Metastases		
TIs <sup>high</sup>	1	-
TIs <sup>intermediate</sup>	1.68 (1.10-2.56)	0.02
TIs <sup>low</sup>	1.81 (1.20-2.74)	0.005
Danish Data		
TIs <sup>high</sup>	1	-
TIs <sup>intermediate</sup>	3.76 (1.23-11.45)	0.02
TIs <sup>low</sup>	2.36 (0.76-7.29)	0.14
Van Allen Data		
TIs <sup>high</sup>	1	-
TIs <sup>intermediate</sup>	3.19 (1.09-9.35)	0.03
TIs <sup>low</sup>	3.50 (1.17-10.51)	0.03
Gide Data		
TIs <sup>high</sup>	1	-
TIs <sup>intermediate</sup>	1.08 (0.33-3.53)	0.9
TIs <sup>low</sup>	4.01 (1.47-10.99)	0.007
Riaz Data		
TIs <sup>high</sup>	1	-
TIs <sup>intermediate</sup>	2.95 (0.92-9.43)	0.07
TIs <sup>low</sup>	4.72 (1.53-14.56)	0.007

Hazard ratios and confidence intervals in the bottom panel correspond to the Kaplan–Meier plots in Fig. 3.

\*Model 6 includes CD8/CD20 groups, stage, type of metastasis, age and gender as covariates.

\*\*Source site corresponds to primary tumour, regional lymph node, regional other and distant metastasis.

---

**Extended Data Table 3 | Immune-related proteins investigated in the GeoMx analysis**

Protein
4.1BB, CD137
ARG1
B7.H3
Bcl.2
Beta.2.microglobulin
CD11c
CD127
CD14
CD163
CD20
CD25
CD27
CD3
CD34
CD4
CD40
CD40L
CD44
CD45
CD45RO
CD56
CD66b
CD68
CD8
CD80
CD86
CTLA4
EpCAM
ER.alpha
FAPalpha
Fibronectin
FOXP3
GAPDH
GITR
GZMB
Her2.ErbB2
Histone.H3
HLA.DR
ICOS
IDO1
Ki.67
LAG3
MART1
Ms.IgG1
Ms.IgG2a
NY.ESO.1
OX40L
PanCk
PD.1
PD.L1
PD.L2
PR
PTEN
Rb.IgG
S100B
S6
SMA
STING
TGFB1
Tim.3
VISTA

Extended Data Table 4 | SAM analysis results to obtain the nine-gene TLS signature

Genes discriminating tumors with CD8 T cells alone compared to immune poor melanomas							Genes discriminating tumors with TLS and CD8 T cells compared to melanomas with CD8 T cells alone	
Gene	Fold change	Gene	Fold change	Gene	Fold change	Gene	Fold change	
C4orf7	5.850525	HLA-DMB	2.1769047	MFNG	1.7937715	HLA-C	1.6125638	
CD8A	4.3367634	ARHGAP9	2.1595395	MSA47	1.7918254	CTSC	1.6079848	
CXCL9	4.282619	LAMP3	2.1572433	TYROBP	1.7895821	RBM47	1.6064883	
CD3D	4.109594	CD3G	2.1555102	ARHGAP25	1.7882159	CFB	1.605289	
NAPSA	3.9003344	EV12B	2.151902	LAPTM5	1.7830758	EV12A	1.6043215	
CXCL13	3.878391	IL18BP	2.1482697	LOC387841	1.7806474	RNASE6	1.6032722	
INDO	3.7693827	FCER1G	2.1476345	EBI3	1.779771	FPR3	1.5993513	
CCL5	3.7572536	CAMK1G	2.1367993	P2RY8	1.7786007	ABI3	1.5974672	
LTB	3.6854622	TBC1D10C	2.131732	C2	1.77573	RGS18	1.5973897	
CCL21	3.6670513	AIM2	2.129987	NCF1	1.774397	LY86	1.5972974	
CCL19	3.634043	CD53	2.111374	PTPN22	1.7705708	PARP9	1.5966645	
IDO1	3.5465145	LCP1	2.1084743	FUCA1	1.7698385	OASL	1.5946562	
GZMK	3.4980235	HLA-F	2.1038625	GIMAP6	1.76921	SERPINA1	1.590805	
CXCL10	3.3749766	FGD3	2.0983796	BTX	1.7660133	SLCO2B1	1.5882876	
GZMA	3.32443	HLA-DRB4	2.096038	OAS2	1.7645061	SUSD3	1.5878158	
LOC652775	3.3141112	BIRC3	2.0951495	OLR1	1.7627058	UBE2L6	1.586905	
UBD	3.2973506	PYHIN1	2.0907834	GBP2	1.7621691	PSME2	1.5857576	
CD69	3.232438	PRF1	2.0893168	TLR8	1.7585747	PLCL2	1.5817555	
NKG7	3.2006474	FAM113B	2.0810633	EBI2	1.7581428	SAMD9L	1.5809402	
CD247	3.1996102	HCLS1	2.078721	RNF126P1	1.7579799	P2RY10	1.5796638	
CD2	3.148034	GVIN1	2.074809	IFIH1	1.7515248	PIK3CG	1.5796467	
GBP5	3.1334765	TNFSF10	2.069712	CCR2	1.7498112	MX1	1.5793523	
GZMB	3.0648828	HLA-DMA	2.0650768	AMICA1	1.7432792	ISG15	1.5779757	
CCR7	3.0349138	RASGRP1	2.0589375	AIF1	1.7415568	CDC42SE2	1.5738891	
CD79A	2.9795752	C1QC	2.056882	GLRX	1.7394656	P2RY13	1.5730076	
CD48	2.9609842	HLA-DPB1	2.0534782	FAM46C	1.7381114	PARVG	1.5702442	
LOC647506	2.8764262	HLA-DOA	2.0447738	MARCO	1.7333692	VAMP5	1.5684881	
HCPC5	2.8615928	PSCDBP	2.0414143	CD86	1.7330478	FOLR2	1.5683817	
SPOCK2	2.8324168	C7	2.0402412	STX11	1.7318233	HLA-DQB2	1.5620369	
PTPRCAP	2.8146172	APOL3	2.0391428	APBB1IP	1.7280052	CD300LF	1.5616324	
LOC649143	2.8042858	PTPN6	2.038657	LYN	1.7253813	VCAM1	1.5614156	
IL32	2.7964118	WAS	2.0366387	TNFAIP8	1.7246032	TNFAIP3	1.5603107	
LAG3	2.7963312	BCL11B	2.0366046	INPP5D	1.724358	SLAMF7	1.5601757	
HLA-DQA1	2.7800684	C20orf100	2.0273168	CYSLTR1	1.7242785	HCK	1.5587567	
SLAMF6	2.7586188	CD37	2.0114977	MCOLN2	1.7232472	CARD11	1.558226	
CD7	2.7440262	PRKCB	2.011462	SOD2	1.718578	CCL3	1.5579741	
LOC100133678	2.7211967	LOC400759	2.0087654	AKNA	1.7166778	PIK3IP1	1.5557806	
HLA-DRB6	2.713655	HLA-H	2.009027	TNFRSF1B	1.7134982	TSTD1	1.555676	
CD27	2.7128735	FGD2	2.0087453	IGSF6	1.7131412	DDX60	1.553378	
RARRES3	2.7052057	SLC40A1	2.0084155	CFD	1.7116679	IFIT2	1.553336	
GBP4	2.6934311	HLA-DRB3	2.0073037	PLCG2	1.7082311	CD72	1.5531305	
LOC651751	2.6899047	NCF1	2.006084	ALOX5	1.7012193	SP140	1.5522988	
LOC652694	2.685632	CECR1	2.0038574	DHR59	1.7011156	ZFP36	1.5516263	
JSRP1	2.6221983	RASAL3	1.9993367	PPP1R16B	1.7008039	UBA7	1.5514268	
GIMAP7	2.6219752	GIMAP5	1.9982749	LAX1	1.6922069	BATF	1.5485026	
IRF8	2.6158779	PIK3AP1	1.9907101	LST1	1.6902115	IFI27	1.547818	
PRKCB1	2.5978892	ITK	1.9889216	ACSL5	1.68995	CSF2RA	1.5462152	
LOC728835	2.573962	CD38	1.9883611	FCGR1B	1.6899358	NCF2	1.5452782	
IGLL3	2.560279	CD96	1.9737343	HLA-E	1.6876732	TAP2	1.5432647	
LOC100133583	2.553927	PTPRC	1.9713553	PIM2	1.681412	DAPP1	1.5431949	
CD52	2.5466175	CCL18	1.9712074	ARHGAP30	1.680812	IFI35	1.5372068	
NAPSB	2.5303907	GIMAP2	1.9646053	HLA-G	1.6804048	ALOX5AP	1.5365016	
SH2D1A	2.5151858	CYBB	1.9622309	SLC7A7	1.6793996	ABCG1	1.5314059	
LOC649923	2.5026162	CXCL11	1.9613086	PSCD4	1.6785691	LYL1	1.5313303	
LOC652493	2.490692	PTGER4	1.9604341	CLEC4A	1.6785455	XAF1	1.5307403	
LOC647450	2.4846182	ITGAL	1.9591348	KYNU	1.678476	RGS10	1.5273899	
CORO1A	2.4811893	MS4A6A	1.952932	CTSH	1.6757919	FLI1	1.526178	
LOC606724	2.4712005	TAP1	1.9535233	IL10RA	1.6756753	IRF7	1.5240406	
HLA-DQB1	2.4548614	GIMAP1	1.943057	HAVCR2	1.6754664	IL4I1	1.5239596	
DNASE1L3	2.4548588	MGC29506	1.9380807	SEPP1	1.6752254	ADRB2	1.5232662	
GZMH	2.4348328	IGLL1	1.9257094	IL15	1.6723912	LAP3	1.5217831	
GBP1	2.4314487	DOCK2	1.9256791	PSTPIP2	1.6706566	APOBEC3G	1.5210005	
HCST	2.4196472	PLA2G7	1.9201738	UCP2	1.667601	TOX2	1.5196278	
PLEK	2.417525	CCL3L3	1.9036595	SIGLEC10	1.666853	LPXN	1.5194254	
FGF2	2.4134262	TRIM22	1.9034454	ZBP1	1.6624216	FBXO6	1.5154316	
CCL13	2.4028354	KLRD1	1.9017122	ITGB2	1.6614282	ASCL2	1.5137706	
IGJ	2.4006696	IFI44L	1.8944559	OAS1	1.6578766	AADACL1	1.5122902	
IRF1	2.3972383	C1QA	1.8862053	LOC647108	1.6575867	MYO1G	1.5105464	
LOC730415	2.377494	IKZF1	1.8854065	IFITM1	1.6569607	IFI6	1.5104185	
CCL4L2	2.3680255	ALDH2	1.8754903	CLECL1	1.65586	EAF2	1.5100641	
GNLY	2.3650563	DOCK8	1.8699276	ICOS	1.653902	LILRB3	1.5093257	
IFNG	2.3623235	SEMA4D	1.8685683	RASSF5	1.6523989	CSF1R	1.5070226	
IL2RB	2.3477395	NCKAP1L	1.859875	C1S	1.651071	CYTH4	1.5061525	
EPSTI1	2.335121	ANKRD22	1.8569533	MS4A4A	1.6509206	MGAT4A	1.505879	
CCL4L1	2.3269346	SLA	1.8484899	LAIR2	1.6478871	LAT2	1.5024157	
LYZ	2.3133543	CTSS	1.8483866	LILRB5	1.646892	SLC31A2	1.5022907	
LOC731682	2.3117633	WARS	1.8480946	LILRB2	1.6428468	ATP8B4	1.5022835	
HLA-DPA1	2.3102183	SASH3	1.8472831	PSMB8	1.642432	RCSN1	1.5018578	
LOC642073	2.3056629	RGS1	1.8432442	CXCR3	1.6415471	GCH1	1.5016657	
HLA-B	2.303435	C1orf162	1.8429196	SPI1	1.6396817	LILRA5	1.5010871	
ADAMDEC1	2.2967832	CRKRS	1.8394536	CXCR4	1.6376555			
STAT1	2.2745712	LRMP	1.8371218	FCGR1A	1.6370385			
CD74	2.273994	SRGN	1.8356284	DEF6	1.6362371			
C1QB	2.2717638	CEBPA	1.8305081	ITGB7	1.6362158			
PVRIG	2.2529182	VNN2	1.8215137	NCF4	1.6340253			
CCL8	2.252525	GIMAP8	1.8203901	AGPAT9	1.6268378			
LOC642113	2.251248	CASP1	1.8195084	LGALS2	1.6255531			
FAM26F	2.244625	DENND2D	1.8194357	CD163	1.6254405			
LIME1	2.2311883	GMFG	1.8193122	FGR	1.62524			
SELL	2.2240884	GPR65	1.8181477	ADD3	1.6250473			
PSMB9	2.2207227	MAP4K1	1.8174882	ANKRD29	1.6237315			
RAC2	2.2163737	ISG20	1.8124862	FCN1	1.623711			
IL7R	2.215939	FBP1	1.8124799	MIR155HG	1.6234839			
FYB	2.2151234	CLEC12A	1.8113737	CD209	1.6229211			
LOC401845	2.2078538	CXCL12	1.8111467	LILRB4	1.6225767			
HLA-DRA	2.1994276	HLA-DOB	1.807083	IFIT3	1.6202953			
KLRB1	2.1980007	STAT4	1.8062018	KIR2DL4	1.6192334			
PLAC8	2.1869752	SLAMF8	1.8041736	TRAF3IP3	1.6186843			
GIMAP4	2.1838875	IFI30	1.7981355	GAPT	1.6185887			
FCRL3	2.1787016	RSAD2	1.7952874	TLR7	1.6158199			
FAM3	2.1784108	PRKCH	1.7944446	STK17B	1.6147563			
CD6	2.1783032	SAMSN1	1.7943485	SLC15A3	1.6138071			

\*Gene unique to the tumours with TLS.

Extended Data Table 5 | Clinical features of the cohorts treated by ICB

	ENTIRE COHORT (N=201)	TLS <sup>HIGH</sup> (N=44)	TLS <sup>INT</sup> (N=57)	TLS <sup>LOW</sup> (N=74)
DANISH ANTI-CTLA4 TREATED COHORT				
COHORT N (%)	37 (18)	13 (35)	12 (32)	12 (32)
AGE AT TREATMENT MEDIAN (RANGE)	63 (33-84)	65 (34-77)	64 (34-80)	59 (33-80)
LDH LEVELS MEDIAN (RANGE)	207 (125-545)	211 (136-302)	184 (147-434)	205 (143-545)
METASTATIC SITE				
SKIN	11	1	5	5
LYMPH NODE	14	6	5	3
CNS	3	1	0	2
LUNG	3	1	1	1
OTHER	3	1	1	1
NA	3	-	-	-
VAN ALLEN ANTI-CTLA4-TREATED COHORT (N=40) <sup>1</sup>				
AGE AT DIAGNOSIS (RANGE)	59 (32-83)	57 (33-83)	61 (43-77)	59 (32-71)
GENDER (%)				
MALE	26 (65)	10 (38)	7 (27)	9 (35)
FEMALE	14 (35)	3 (21)	6 (43)	5 (36)
RIAZ ANTI-PD1 TREATED COHORT (N=40) <sup>2</sup>				
PRIOR ANTI-CTLA4 TREATMENT (%)				
YES	19 (48)	8 (42)	7 (37)	4 (21)
NO	21 (52)	5 (24)	6 (29)	10 (48)
M STAGE				
M1A	10	4	4	2
M1B	7	4	1	2
M1C	16	4	5	7
NA	6	-	3	3
GIDE ANTI-PD1 TREATED COHORT (N=61) <sup>2</sup>				
GENDER N (%)				
MALE	45 (65)	14 (31)	16 (36)	15 (33)
FEMALE	24 (35)	9 (38)	7 (29)	8 (33)
AGE MEDIAN (RANGE)	61 (24-90)	68 (24-78)	62 (40-82)	58 (42-90)
METASTATIC SITE				
LYMPH NODE	19	9	6	4
SUBCUTANEOUS	43	10	17	16
LUNG	3	3	-	-
OTHER	4	1	-	3

## Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- | n/a                                 | Confirmed  |
|-------------------------------------|--|
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided<br><i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>   |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> A description of all covariates tested   |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> For null hypothesis testing, the test statistic (e.g. $F$ , $t$ , $r$ ) with confidence intervals, effect sizes, degrees of freedom and $P$ value noted<br><i>Give <math>P</math> values as exact values whenever suitable.</i>                            |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Estimates of effect sizes (e.g. Cohen's $d$ , Pearson's $r$ ), indicating how they were calculated  |

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

#### Data collection

The following software was used in the study:

R version 3.5.1 with base/standard packages and additional packages: estimate\_1.0.13, GLAD\_2.46.0, limma\_3.38.3, survival\_2.42-3, swamp\_1.4.1 and TxDb.Hsapiens.UCSC.hg19.knownGene\_3.2.2 . SAM analysis in the TMeV tool (v 4.8.1). HISAT (v. 2.1.0), Stringtie (v. 1.3.3b) (<https://ccb.jhu.edu/software/stringtie/>). samtools 1.9 (<https://sourceforge.net/projects/samtools/files/samtools/>). CONTRA 2.0.3. (<http://contra-cnv.sourceforge.net/>)

#### Data analysis

The following software was used in the study:

R version 3.5.1 with base/standard packages and additional packages: estimate\_1.0.13, GLAD\_2.46.0, limma\_3.38.3, survival\_2.42-3, swamp\_1.4.1 and TxDb.Hsapiens.UCSC.hg19.knownGene\_3.2.2 . SAM analysis in the TMeV tool (v 4.8.1). HISAT (v. 2.1.0), Stringtie (v. 1.3.3b) (<https://ccb.jhu.edu/software/stringtie/>). samtools 1.9 (<https://sourceforge.net/projects/samtools/files/samtools/>). CONTRA 2.0.3. (<http://contra-cnv.sourceforge.net/>)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.



## Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All relevant data are available and are included with the manuscript as source data. Digital spatial profiling data used in Fig. 2 and gene expression microarray data from anti-CTLA4 treated patients are available as source data. Data from public repositories were accessed from GSE65904 (Cirenajwis et al.), TCGA data portal SKCM level 3 release 3.1.14.0 (TCGA data), PRJEB23709 (Gide et al.), [https://github.com/riazn/bms038\\_analysis/tree/master/data](https://github.com/riazn/bms038_analysis/tree/master/data) (Riaz et al.), GSE115978 (Jerby-Arnon et al.) and GSE120575 (Sade-Feldman et al.). Any other relevant data and code can be obtained from the corresponding authors upon reasonable request.

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	These cohorts are unique and sample sizes are for the Lund cohort (n=177) and the Danish anti-CTLA4 (n=37). These are cohorts collected during a long time and we believe these are sufficient to ask the questions we put forward in the manuscript. In addition, we use cohorts where data are available in public repositories.
Data exclusions	Patients were only excluded from analyses if the tissue specimen did not include any tumor cells as determined by a dermatopathologist
Replication	The derived TLS gene expression signature was firmly validated in several different datasets across multiple gene expression platforms.
Randomization	Randomization is not relevant for this study as we are analysing tumor specific phenotypes.
Blinding	Immunostaining evaluation were done blinded by three independent readers of which one was a board-certified dermatopathologist.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input type="checkbox"/>	<input checked="" type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data

### Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

## Antibodies

Antibodies used	The following antibodies are used in this study. CD3 (polyclonal, cat no A0452, lot no 20066809, Dako/Agilent), CD8 (clone C8/144B, cat no M7103, lot no 20029542, Dako/agilent), CD20 (clone L26, cat no. 760-2531, lot no 00064779, Dako/Agilent), SOX10 (clone BC34, cat no. ACI 3099 A, C, Biocare), B2M (polyclonal, cat no. A0072, lot no 00066626, Agilent/Dako), Ki67 (clone MIB-1, cat no. GA626, lot no 20027876, Dako/agilent) . For immunofluorescence rabbit-anti-CXCR5 (cat no. 3180237-9, lot no GR3229212-1, abcam) and rabbit-anti-CXCL13 (cat no NBP2-1604155, lot no 0141712Da843058, Novus Biologicals)
Validation	Antibodies used in this study are well validated. This is clearly demonstrated by the manufacturer as well as the vast number of citations refereing to the antibodies. In addition, the SOX10 and CD20 stainings were performed at a Swedish clinically approved diagnostic laboratory. The following information is available for each antibody. CD3 (polyclonal, cat no A0452) - Rabbit Anti-Human for IHC - In Western blotting, the antibody detects bands of the expected

molecular weights for CD3 antigens. The antibody recognizes CD3e in both a T-cell line (Jurkat) and a natural killer cell line (NK11), but does not react with lysates prepared from several B-cell lines (Raji, Ramos and JY), a myeloid cell line (U937) or a colon carcinoma cell line (Colo-205).

In immunoprecipitation from Nonidet P40 lysates of surface-iodated T lymphoblasts, the antibody precipitates the g (26 kDa), d (21 kDa) and e (19 kDa) chain of the CD3 molecule, similar to the precipitation pattern seen using the well-characterized monoclonal mouse anti-human CD3, clone UCHT1.

In ELISA, the antibody labels the CD3 peptide used as immunogen.

CD8 (clone C8/144B, cat no M7103) - Mouse Anti-Human - SDS-PAGE analysis of immunoprecipitates formed between lysates of 125I-labeled human T lymphoblasts and the antibody shows reaction primarily with a 32 kDa polypeptide corresponding to CD8 $\alpha$ .

CD20 (clone L26, cat no. 760-2531) - Mouse Anti-Human - The antibody was clustered as anti-CD20 at the Fifth International Workshop and Conference on Human Leucocyte Differentiation Antigens held in Boston 1993. SDS-PAGE analysis of immunoprecipitates formed between 125I-labeled tonsil cell lysate and the antibody shows reaction primarily with 30 kDa and 33 kDa polypeptides. Studies using COS-1 cells transfected with cDNA encoding the CD20 molecule, indicate that the antibody labels an intracytoplasmic epitope localized on the CD20 molecule.

SOX10 (clone BC34, cat no. ACI 3099 A, C, Biocare) - Mouse Anti-Human - Nuclear staining of SOX10 [BC34] was observed in 96.4% (106/110) of cases of cutaneous melanoma and 83.9% (73/87) of cases of metastatic melanoma (Table 1). Staining of SOX10 [BC34] was also observed in spindle cell melanoma (100%, 19/19), desmoplastic melanoma (96.6%, 28/29), benign nevi (100%, 20/20) and schwannomas (100%, 28/28). SOX10 [BC34] nuclear staining was observed in the expected normal tissues: oligodendrocytes in cerebrum and cerebellum, myoepithelial cells in breast and salivary glands, melanocytes in skin, and Schwann cells in peripheral nerve.

B2M (polyclonal, cat no. A0072) - rabbit anti-human - Crossed immunoelectrophoresis: B2M precipitation curve is visible in the usage of 12.5  $\mu$ l A0072 per cm<sup>2</sup> gel surface against concentrated urine from patients with tubular proteinuria. Using 2  $\mu$ l human plasma no precipitation is observed. Staining: Coomassie Brilliant Blue.

Ki67 (clone MIB-1) - Mouse Anti-Human - In Western blotting of lysates of the multiple myeloma cell line, IM-9, the MIB-1 antibody labels bands of 345 and 395 kDa, identical to the bands labeled by the original Ki-67 antibody. Furthermore, Western blotting and competitive binding experiments clearly demonstrate that MIB-1, like the original Ki-67 antibody, reacts with an epitope encoded by a 66 bp repetitive element in the Ki-67 gene. In immunohistochemistry, the MIB-1 and the Ki-67 antibodies provide identical staining patterns on serial tonsillar frozen sections. The MIB-1 antibody recognizes native Ki-67 antigen and recombinant fragments of the Ki-67 molecule.

IF-specific antibodies

rabbit-anti-CXCR5 (cat no. 3180237-9) - rabbit anti-human - In Western Blot one specific band is observed in mouse B cells. Additional evidence comes from IF experiments on lymphoma cells.

rabbit-anti-CXCL13 (cat no NBP2-1604155) - rabbit anti-human - In Western Blot one specific band is observed in 293T whole extract. Additional evidence comes from IF experiments on HepG2 cells.

## Human research participants

Policy information about [studies involving human research participants](#)

Population characteristics	Overall, 104 patients had regional metastatic disease, 50 distant disease and 19 local disease. Four patients were of unknown stage. This is a historic cohort collected between 2000 and 2012. 57% were male patients, the age at diagnosis were on average 65. 64% were lymph node metastases, 20% were subcutaneous metastases, 6% were visceral metastases and 8% were primary tumors.
Recruitment	The Lund pre-checkpoint inhibitor cohort was collected prospectively from 1997-2012 at the Dept of Surgery, Skåne University Hospital in Sweden. Only patients where sufficient tissue was available were included in the study. The Danish anti-CTLA4 treated cohort is retrospective and includes all patients receiving first-line anti-CTLA4 treatment up until year 2016.
Ethics oversight	This study was approved by the Regional Ethics Committee at Lund University (Dnr. 191/2007 and 101/2013). The sample cohort, representing a population-based retrospective collection (n=177), was obtained at the Department of Surgery at Skåne University Hospital. We also collected paraffin embedded tumor tissue from 119 patients of which 37 patients received anti-CTLA4 as first-line therapy. These patients were collected in Denmark and available biopsy was obtained a maximum of six months before therapy start. This study was approved by the regional ethical committee (H-15010200).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

# TGF- $\beta$ orchestrates fibrogenic and developmental EMTs via the RAS effector RREB1

<https://doi.org/10.1038/s41586-019-1897-5>

Received: 4 October 2018

Accepted: 5 November 2019

Published online: 8 January 2020

There are amendments to this paper

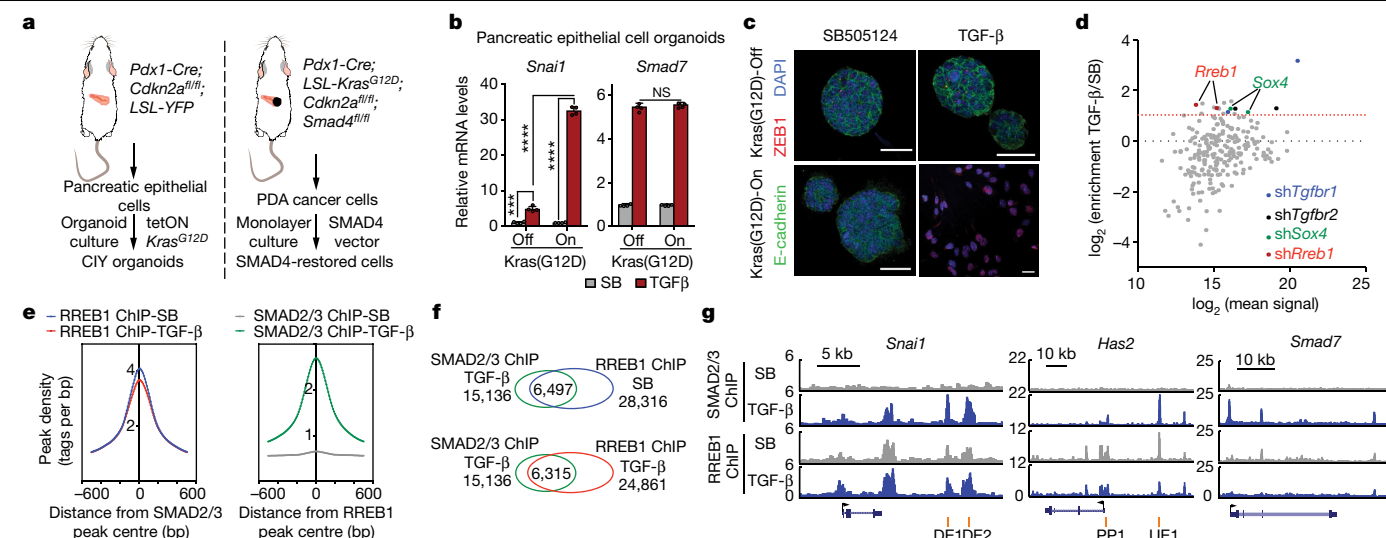
Jie Su<sup>1</sup>, Sophie M. Morgani<sup>2,3</sup>, Charles J. David<sup>1,6</sup>, Qiong Wang<sup>1,7</sup>, Ekrem Emrah Er<sup>1</sup>, Yun-Han Huang<sup>1,4</sup>, Harihar Basnet<sup>1</sup>, Yilong Zou<sup>1,4,8</sup>, Weiping Shu<sup>1</sup>, Rajesh K. Soni<sup>5</sup>, Ronald C. Hendrickson<sup>5</sup>, Anna-Katerina Hadjantonakis<sup>2</sup> & Joan Massagué<sup>1\*</sup>

Epithelial-to-mesenchymal transitions (EMTs) are phenotypic plasticity processes that confer migratory and invasive properties to epithelial cells during development, wound-healing, fibrosis and cancer<sup>1–4</sup>. EMTs are driven by SNAIL, ZEB and TWIST transcription factors<sup>5,6</sup> together with microRNAs that balance this regulatory network<sup>7,8</sup>. Transforming growth factor  $\beta$  (TGF- $\beta$ ) is a potent inducer of developmental and fibrogenic EMTs<sup>4,9,10</sup>. Aberrant TGF- $\beta$  signalling and EMT are implicated in the pathogenesis of renal fibrosis, alcoholic liver disease, non-alcoholic steatohepatitis, pulmonary fibrosis and cancer<sup>4,11</sup>. TGF- $\beta$  depends on RAS and mitogen-activated protein kinase (MAPK) pathway inputs for the induction of EMTs<sup>12–19</sup>. Here we show how these signals coordinately trigger EMTs and integrate them with broader pathophysiological processes. We identify RAS-responsive element binding protein 1 (RREB1), a RAS transcriptional effector<sup>20,21</sup>, as a key partner of TGF- $\beta$ -activated SMAD transcription factors in EMT. MAPK-activated RREB1 recruits TGF- $\beta$ -activated SMAD factors to *SNAIL*. Context-dependent chromatin accessibility dictates the ability of RREB1 and SMAD to activate additional genes that determine the nature of the resulting EMT. In carcinoma cells, TGF- $\beta$ –SMAD and RREB1 directly drive expression of SNAIL and fibrogenic factors stimulating myofibroblasts, promoting intratumoral fibrosis and supporting tumour growth. In mouse epiblast progenitors, Nodal–SMAD and RREB1 combine to induce expression of SNAIL and mesendoderm-differentiation genes that drive gastrulation. Thus, RREB1 provides a molecular link between RAS and TGF- $\beta$  pathways for coordinated induction of developmental and fibrogenic EMTs. These insights increase our understanding of the regulation of epithelial plasticity and its pathophysiological consequences in development, fibrosis and cancer.

Oncogenic mutations in *KRAS* are prevalent in pancreatic adenocarcinoma (PDA) and strongly potentiate the induction of EMT by TGF- $\beta$ <sup>12</sup>. We transduced an inducible *KRAS*(G12D) oncogene into pancreatic epithelial organoids from *Pdx1-cre;Cdkn2a<sup>fl/fl</sup>;lox-stop-lox* (*LSL*)-YFP (CIY) mice (Fig. 1a), and treated the organoids with either TGF- $\beta$  or SB505124<sup>22</sup> (SB), which blocks endogenous TGF- $\beta$  signalling. Before induction of *KRAS*(G12D) expression, TGF- $\beta$  caused a modest (fourfold) increase in *Snail* expression and did not alter organoid morphology or survival. When *KRAS*(G12D) was induced, TGF- $\beta$  treatment caused a 30-fold increase in *Snail* expression (Fig. 1b), followed by a decrease in E-cadherin, increase in ZEB1, organoid dissociation (Fig. 1c, Extended

Data Fig. 1a) and apoptosis (Supplementary Video 1), all characteristic of a lethal EMT<sup>12</sup>. Induction of *Smad7* expression, a conserved TGF- $\beta$  negative-feedback response, was independent of *KRAS*(G12D) (Fig. 1b). TGF- $\beta$  modulated the expression of 56 genes by more than fourfold and *KRAS*(G12D) augmented TGF- $\beta$  induction of 13 of these genes (Extended Data Fig. 1b, c), including *Snail* and hyaluronan synthase 2 (*Has2*), known regulators of EMT<sup>23</sup> (Extended Data Fig. 1d). We confirmed this response pattern in different pancreatic organoids and primary cultures (Extended Data Fig. 1e, f). These TGF- $\beta$  responses required SMAD4, as shown in PDA cells with restored SMAD4 expression (Extended Data Fig. 1g).

<sup>1</sup>Cancer Biology and Genetics Program, Sloan Kettering Institute, Memorial Sloan Kettering Cancer Center, New York, NY, USA. <sup>2</sup>Developmental Biology Program, Sloan Kettering Institute, Memorial Sloan Kettering Cancer Center, New York, NY, USA. <sup>3</sup>Wellcome Trust–Medical Research Council Centre for Stem Cell Research, University of Cambridge, Cambridge, UK. <sup>4</sup>Gerstner Sloan Kettering Graduate School of Biomedical Sciences, Memorial Sloan Kettering Cancer Center, New York, NY, USA. <sup>5</sup>Microchemistry and Proteomics, Memorial Sloan Kettering Cancer Center, New York, NY, USA. <sup>6</sup>Present address: Tsinghua University School of Medicine, Department of Basic Sciences, Beijing, China. <sup>7</sup>Present address: Department of Histo-embryology, Genetics and Developmental Biology, Shanghai Key Laboratory of Reproductive Medicine, Shanghai Jiao Tong University School of Medicine, Shanghai, China. <sup>8</sup>Present address: Chemical Biology and Therapeutics Science program, Broad Institute, Cambridge, MA, USA. \*e-mail: j-massague@ski.mskcc.org



**Fig. 1 | RREB1 is a KRAS-dependent SMAD cofactor.** **a**, Source and generation of CIY pancreatic epithelial organoids and SMAD4-restored PDA cells. **b**, *Snai1* and *Smad7* mRNA levels in pancreatic epithelial organoid cultures. Cells engineered to express doxycycline-inducible KRAS(G12D) treated with the TGF- $\beta$  and Nodal receptor inhibitor SB505124 (SB, 2.5  $\mu$ M) or TGF- $\beta$  (10 pM) for 1.5 h. Data are mean  $\pm$  s.d.;  $n = 4$ ; two-way analysis of variance (ANOVA), \*\*\*\* $P < 0.0001$ . **c**, E-cadherin, ZEB1 and DAPI immunofluorescence in CIY pancreatic organoids with or without KRAS(G12D) expression, treated with SB505124 or TGF- $\beta$  for 2.5 days. Scale bars, 30  $\mu$ m. Images are representative of two independent experiments. **d**, Screening of pancreatic progenitor transcription factor shRNA library for mediators of TGF- $\beta$ -induced lethal EMT. Dot plot of shRNA enrichment in TGF- $\beta$ -treated versus SB505124-treated SMAD4-restored PDA cells. *Sox4*<sup>12</sup> and *Rreb1* transcription factor genes score

## RREB1 as a RAS-regulated SMAD cofactor

TGF- $\beta$  binding to the receptor kinases TGFBR1 and TGFBR2 activates SMAD2–SMAD3–SMAD4 (SMAD2/3/4) trimeric complexes, which target specific promoters and enhancers by interacting with context-determining transcription factors<sup>9</sup>. SMAD2/3 chromatin immunoprecipitation and DNA sequencing (ChIP-seq) in PDA cells treated with TGF- $\beta$  revealed binding motifs for various RAS transcriptional effectors (FOS and JUN AP-1 components and ELK3 and the SMAD binding motifs CAGAC and 5GC<sup>24</sup> within SMAD2/3 peaks independently of KRAS(G12D) (Extended Data Fig. 1h–j). Notably, RREB1 motifs were specifically enriched within SMAD2/3 peaks in KRAS(G12D)-dependent TGF- $\beta$  targets (Extended Data Fig. 1h). Although EMT is generally pro-tumorigenic in carcinoma cells, TGF- $\beta$  triggers apoptosis in KRAS-mutant pancreatic progenitors owing to simultaneous induction of SNAIL and the pro-epithelial transcription factor SOX4<sup>12</sup>. We used this property of KRAS-mutant pancreatic progenitors to screen a shRNA library targeting 40 transcription factors expressed in PDA cells using shRNAs targeting the TGF- $\beta$  receptors as positive controls (Fig. 1d). *Rreb1* and *Sox4* were the only transcription factor transcripts for which two independent shRNAs were enriched more than twofold (Fig. 1d).

RREB1 contains 15 zinc fingers<sup>21</sup>, but little is known about its function and regulation<sup>25–27</sup>. In PDA cells in which SMAD4 expression has been restored and that also express haemagglutinin (HA)-tagged RREB1 (residues 1–1291 mouse isoform) (Extended Data Fig. 2a), ligation assays showed close proximity between nuclear RREB1 and SMAD2/3 following TGF- $\beta$  treatment (Extended Data Fig. 2b, c). Co-immunoprecipitation revealed interactions between SMAD3 and HA–RREB1 (Extended Data Fig. 2d). The genome-binding pattern of HA–RREB1 overlapped with that of SMAD2/3 in TGF- $\beta$ -treated cells (Fig. 1e, f, Extended Data Fig. 2e),

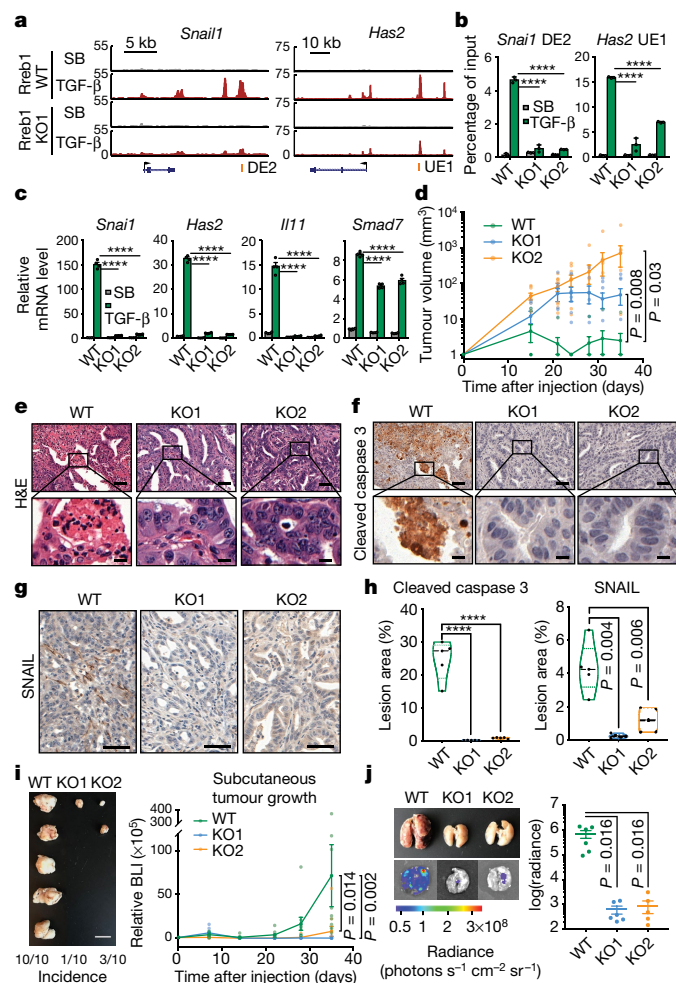
positive in the screen. shRNAs targeting *Tgfr1* and *Tgfr2* are included as positive controls. **e**, Position of RREB1 peak summits relative to summits of overlapping SMAD2/3 peaks (left), and position of SMAD2/3 peak summits relative to summits of overlapping RREB1 peaks (right), based on ChIP-seq analysis (Extended Data Fig. 2e). **f**, Venn diagram showing overlap between SMAD2/3 and RREB1 ChIP-seq peaks, based on ChIP-seq analysis in Extended Data Fig. 2e. **g**, Gene track view of SMAD2/3 and HA–RREB1 ChIP-seq tags at indicated loci and experimental conditions. Gene bodies are represented below the track sets. PP, proximal promoter; DE, downstream enhancer; UE, upstream enhancer. ChIP-seq was performed once and an independent ChIP was performed in which selective genomic regions were confirmed by quantitative PCR (qPCR). See also Extended Data Figs. 1–3 and Supplementary Video 1.

including in *Snai1* and *Has2* but not in *Smad7* (Fig. 1g). HA–RREB1 bound to these loci in the absence of TGF- $\beta$  signalling (Fig. 1e–g, Extended Data Fig. 2e). MAPK signalling has previously been implicated in RREB1 regulation<sup>28</sup>. Treatment of SMAD4-restored PDA cells with the ERK inhibitor SCH772984 (ERKi) or the MEK inhibitor AZD6244 (MEKi) did not alter nuclear localization (Extended Data Fig. 3a) or levels of RREB1 (Extended Data Fig. 3b, c), but diminished binding of HA–RREB1 to *Snai1*, *Has2* and *Il11* cis-regulatory regions (Extended Data Fig. 3d). HA–RREB1 immunoprecipitated from PDA cell lysates bound double-stranded DNA probes corresponding to *Snai1* enhancer and *Has2* promoter regions; ERKi treatment decreased this activity (Extended Data Fig. 3e). We identified four ERK-dependent phosphorylation sites in HA–RREB1 immunoprecipitated from SMAD4-restored PDA cells (Extended Data Fig. 3f, g); all were situated between zinc-finger domains (Extended Data Fig. 3h). S161 and S970 fit the MAPK phosphorylation motif PX(S/T)P, whereas S1138 and S175 may represent indirect phosphorylation by other kinases. RREB1 with S161 or S970 alanine substitutions was deficient in restoring *Snai1* and *Has2* TGF- $\beta$  responses to *Rreb1*-knockout cells and in binding to these loci, compared with vectors encoding RREB1 with phosphorylation-mimicking aspartate substitution (Extended Data Fig. 3i, j).

## RREB1 and TGF- $\beta$ -dependent EMT

*Rreb1* knockout in SMAD4-restored PDA cells (Extended Data Fig. 4a–c) reduced TGF- $\beta$ -dependent binding of SMAD2/3 to regulatory regions in *Snai1* and *Has2*, and abolished their induction and EMT (Fig. 2a–c, Extended Data Fig. 4d, e). *Rreb1* knockout had limited effects on the binding of SMAD2/3 to, and induction of, *Smad7* (Fig. 2c, Extended Data Fig. 4f). Restoration of RREB1 rescued induction of *Snai1*, *Has2* and *Il11* by TGF- $\beta$  in *Rreb1*-knockout cell lines (Extended Data Fig. 4g).





**Fig. 2 | RREB1 mediates KRAS and TGF-β dependent EMT.** **a**, Gene track view of SMAD2/3 ChIP-seq tags at indicated loci of *Rreb1* wild type (WT) and *Rreb1*-knockout (KO) SMAD4-restored mouse PDA cells. ChIP-seq performed once and confirmed for selected genomic regions by ChIP-PCR. **b**, ChIP-PCR analysis of SMAD2/3 binding to indicated sites of *Snail1* (DE2) and *Has2* (UE1) in WT and *Rreb1*-KO PDA cells after treatment with SB505124 (2.5 μM) or TGF-β (100 pM) for 1.5 h. Data are mean ± s.e.m.; *n* = 4; two-way ANOVA; \*\*\*\**P* < 0.0001. **c**, Levels of *Snail1*, *Has2*, *Il11* and *Smad7* in WT and *Rreb1*-KO PDA cells after treatment with SB505124 or TGF-β for 1.5 h. Data are mean ± s.e.m.; *n* = 4; two-way ANOVA; \*\*\*\**P* < 0.0001. **d**, Volume of WT and *Rreb1*-KO SMAD4-restored PDA subcutaneous tumours in syngeneic mice. Data are mean ± s.e.m.; *n* = 10 tumours, 5 mice per group; two-way ANOVA. **e–g**, Representative haematoxylin and eosin (H&E) staining (**e**), cleaved caspase-3 immunohistochemistry (**f**) and SNAIL immunohistochemistry (**g**) of subcutaneous tumours formed by WT and *Rreb1*-KO SMAD4-restored PDA cells 35 d after inoculation. Scale bars (**e**, **f**, top), 50 μm; (**e**, **f**, bottom), 10 μm; (**g**) 50 μm. In **e–g**, Images are representative of five biological replicates. **h**, Quantification of cleaved caspase-3-positive and SNAIL-positive cells in PDA tumour sections. *n* = 5 per group; two-tailed unpaired *t*-test; \*\*\*\**P* < 0.0001. In violin plots, the middle line shows the median and dotted lines represent first and third quartiles. **i**, Left, subcutaneous tumours formed by WT or *Rreb1*-KO 393T3 lung adenocarcinoma cells in syngeneic B6129Sf1/J mice excised 35 days after inoculation. Scale bar, 10 mm. Right, tumour growth monitored by firefly luciferase bioluminescence imaging (BLI) plotted over time. Data are mean ± s.e.m.; *n* = 10 tumours, 5 mice per group; two-way ANOVA. **j**, Representative ex vivo lung bright-field and BLI images from mice inoculated 21 d after via tail-vein inoculation of WT or *Rreb1*-KO 393T3 cells. Lung colonization load was quantified by BLI. Data are mean ± s.e.m.; *n* = 6 mice per group; two-tailed unpaired *t*-test. See also Extended Data Figs. 4–6.

The induction of lethal EMT by TGF-β in KRAS-mutant pancreatic progenitor cells is a barrier to PDA development<sup>12</sup>. SMAD4-restored PDA cells grew poorly as subcutaneous tumours in mice (Fig. 2d), were

undifferentiated (Fig. 2e) and contained cells expressing apoptosis markers (Fig. 2f, h) and SNAIL (Fig. 2g, h). By contrast, *Rreb1*-knockout cells had higher tumorigenic activity (Fig. 2d), with well-differentiated epithelial histology (Fig. 2e) and few apoptotic (Fig. 2f, h) or SNAIL<sup>+</sup> cells (Fig. 2g, h). Notably, *RREB1* is downregulated in human PDA<sup>25</sup> and mutated in approximately 5% of PDA cases<sup>20</sup>.

Activating KRAS mutations define a major subtype of human lung adenocarcinoma (LUAD). 393T3 cells derived from a *Kras*<sup>G12D</sup>; *p53*<sup>-/-</sup> mouse LUAD tumour<sup>29</sup> showed ERK-dependent induction of *Snail1* and *Has2* by TGF-β, followed by EMT without apoptosis (Extended Data Fig. 5a–e). *Rreb1* knockout inhibited the induction of EMT by TGF-β and acutely diminished growth of 393T3 cells as subcutaneous tumours and pulmonary metastatic colonies in mice (Fig. 2i, j, Extended Data Fig. 5f–j). In A549, a KRAS-mutant human LUAD cell line<sup>30</sup>, *RREB1* knockout (Extended Data Fig. 6a) diminished *SNAIL1*, *SNAIL2* (which encodes SLUG) and EMT responses to TGF-β, and inhibited tumour formation in mice (Extended Data Fig. 6b–d). Collectively, the results indicate that RREB1 mediates TGF-β-induced EMT in PDA and LUAD models independently of the tumorigenic phenotype associated with EMT.

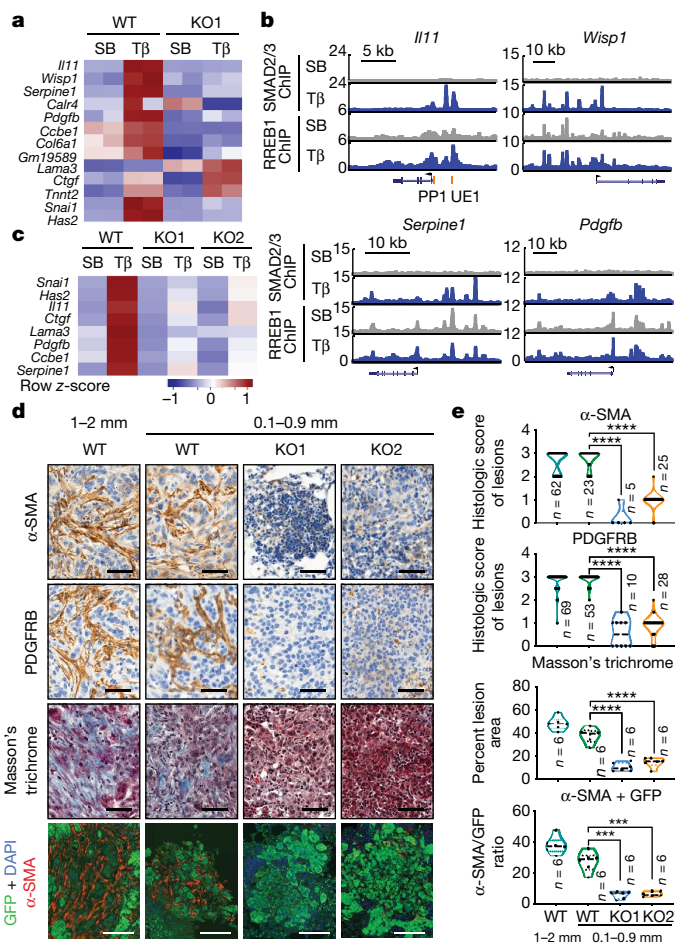
## EMT-associated fibrogenic program

The KRAS-dependent TGF-β response in pancreatic cancer progenitors showed enrichment for gene signatures of cell adhesion, migration and EMT (Extended Data Fig. 6e). Notably, a majority of the 13 KRAS-dependent genes induced by TGF-β were related to deposition of fibrous connective tissue (Extended Data Fig. 1d). Four of these genes encode inducers of extracellular matrix (ECM) production by mesenchymal cells in fibrosis, including interleukin 11 (IL-11) in cardiovascular and renal fibrosis<sup>31</sup>, connective tissue growth factor (CTGF, also known as CCN2) in glomerulonephritis<sup>32</sup>, WNT-inducible signalling pathway protein 1 (WISP1, also known as CCN4) in idiopathic pulmonary fibrosis<sup>33</sup>, and platelet-derived growth factor B (PDGFB) in hepatic fibrosis<sup>34</sup>. The gene set additionally includes the ECM proteins laminin-α3 (*Lama3*), collagen 6a1 (*Col6a1*), collagen and calcium-binding EGF domain-containing protein 1 (*Ccbe1*), the ECM protease inhibitor serpin E1 (*Serpine1*) and *Has2*.

Induction of *Il11*, *Wisp1*, *Serpine1*, *Pdgfb*, *Ccbe1*, *Has2* and *Col6a1* by TGF-β in mouse PDA cells required RREB1 (Fig. 3a, Extended Data Fig. 6f). RREB1 ChIP peaks overlapped with SMAD2/3 peaks in these genes (Fig. 3b). In PDA cells, TGF-β induced expression of *Snail1* and *Zeb1* as previously described<sup>35</sup> (Extended Data Fig. 6g), and depletion of SNAIL and ZEB1 (Extended Data Fig. 6h, i) inhibited EMT but not fibrogenic gene responses (Extended Data Fig. 6j–m), showing that these gene responses are integral, but experimentally divisible, components of a common fibrogenic EMT program. Similar RREB1-dependent induction of these fibrogenic genes and *Snail1* by TGF-β occurred in 393T3 and A549 LUAD cells (Fig. 3a, c). 393T3 pulmonary nodules showed marked presence of cancer-associated myofibroblasts and abundant collagen deposition, whereas time-matched, size-matched *Rreb1*-knockout 393T3 nodules did not (Fig. 3d, e). Thus, TGF-β-activated SMADs converge with RAS-activated RREB1 to drive fibrogenic EMTs in PDA and LUAD cells.

Mammary ductal morphogenesis involves EMT<sup>36</sup>. Mammary epithelial cells undergo EMT in response to TGF-β<sup>37</sup>; EMT induction by TGF-β in normal mouse mammary gland (NMuMG) cells<sup>38,39</sup> requires ERK<sup>40</sup> and RREB1 (Extended Data Fig. 7a–c). RREB1 mediated SMAD2/3 binding to the *Snail1* locus and, to a lesser extent, the *Has2* locus, and induction of these genes by TGF-β (Extended Data Fig. 7d–f). ERKi diminished binding of HA-RREB1 to regulatory regions of *Snail1* and *Has2* in NMuMG cells (Extended Data Fig. 7g). The ERK-pathway activator epidermal growth factor (EGF) increased—and ERKi suppressed—these gene responses, whereas an inhibitor of the EGF receptor had little effect on basal *Snail1* and *Has2* expression (Extended Data Fig. 7h),





**Fig. 3 | RREB1 mediates a TGF- $\beta$  fibrogenic response.** **a**, Heat map of fibrogenic gene responses in WT and *Rreb1*-KO PDA cells after treatment with SB505124 or TGF- $\beta$  (T $\beta$ ) for 1.5 h. *n* = 2. **b**, Gene track view of SMAD2/3 and HA-RREB1 ChIP-seq tags at indicated loci and experimental conditions. Gene bodies represented at bottom of track sets. ChIP-seq was performed once and an independent ChIP was performed in which selective genomic regions were confirmed by qPCR. **c**, Heat map of fibrogenic genes in WT and *Rreb1*-KO 393T3 cells treated with SB505124 or TGF- $\beta$  for 1.5 h. *n* = 4. **d**, Representative  $\alpha$ -smooth muscle actin ( $\alpha$ -SMA) and PDGFRB immunohistochemistry, Masson's trichrome stain and  $\alpha$ -SMA and GFP immunofluorescence of colonized lung tissue after tail vein injection of WT or *Rreb1*-KO GFP<sup>+</sup> 393T3 cells. Scale bars, 100  $\mu$ m. **e**, Quantification of staining in **d**. *n* for each group indicated in graph: two-tailed unpaired *t*-test; \*\*\*\**P* < 0.0001, \*\*\**P* < 0.001. Violin plots show all data points, midline represents the median and dotted lines show first and third quartiles. See also Extended Data Figs. 6, 7.

indicating that RREB1 is required for TGF- $\beta$ -induced EMT in normal mammary epithelial cells.

### Basis for contextual EMT programs

Next, we investigated the role of RREB1 during gastrulation, whereby pluripotent epiblast cells undergo an EMT as they migrate and differentiate. Nodal, FGF and WNT signals drive mesendodermal differentiation and EMT in epiblast cells<sup>1,3,4</sup>. In a spatially resolved RNA-sequencing (RNA-seq) dataset<sup>41</sup>, *Rreb1* transcripts accumulated in the posterior primitive streak domain at mid-gastrulation (embryonic day (E)7.0) (Extended Data Fig. 8a), and overlapped with mesendoderm markers *Gsc* and *Brachyury* (also known as *T*) and EMT markers *Snai1* and *Cdh2* (N-cadherin) (Extended Data Fig. 8a). Mouse embryonic stem (ES) cells form embryoid bodies recapitulating signalling and lineage specification events of gastrulation<sup>42</sup>. Expression of the mesendoderm

genes *Eomes*, *Mixl1*, *T*, goosecoid (also known as *Gsc*), *Fgf8* and *Wnt3* gradually increased after two days of embryoid body differentiation, peaking on day 4 together with EMT drivers *Snai1*, *Twist1*, *Twist2* and *Zeb2*, and *Cdh2* (Fig. 4a, Extended Data Fig. 8b). EMT, stem cell differentiation and gastrulation transcriptional signatures were enriched in parallel (Fig. 4b). *Rreb1* knockout (Extended Data Fig. 8c) inhibited the expression of *Snai1* and key mesendoderm genes (Extended Data Fig. 8d). Addition of activin A (ligand for Nodal receptors) to day 3 embryoid bodies augmented the expression of mesendoderm and *Snai1* genes in wild-type but not *Rreb1*-knockout embryoid bodies (Extended Data Fig. 8e).

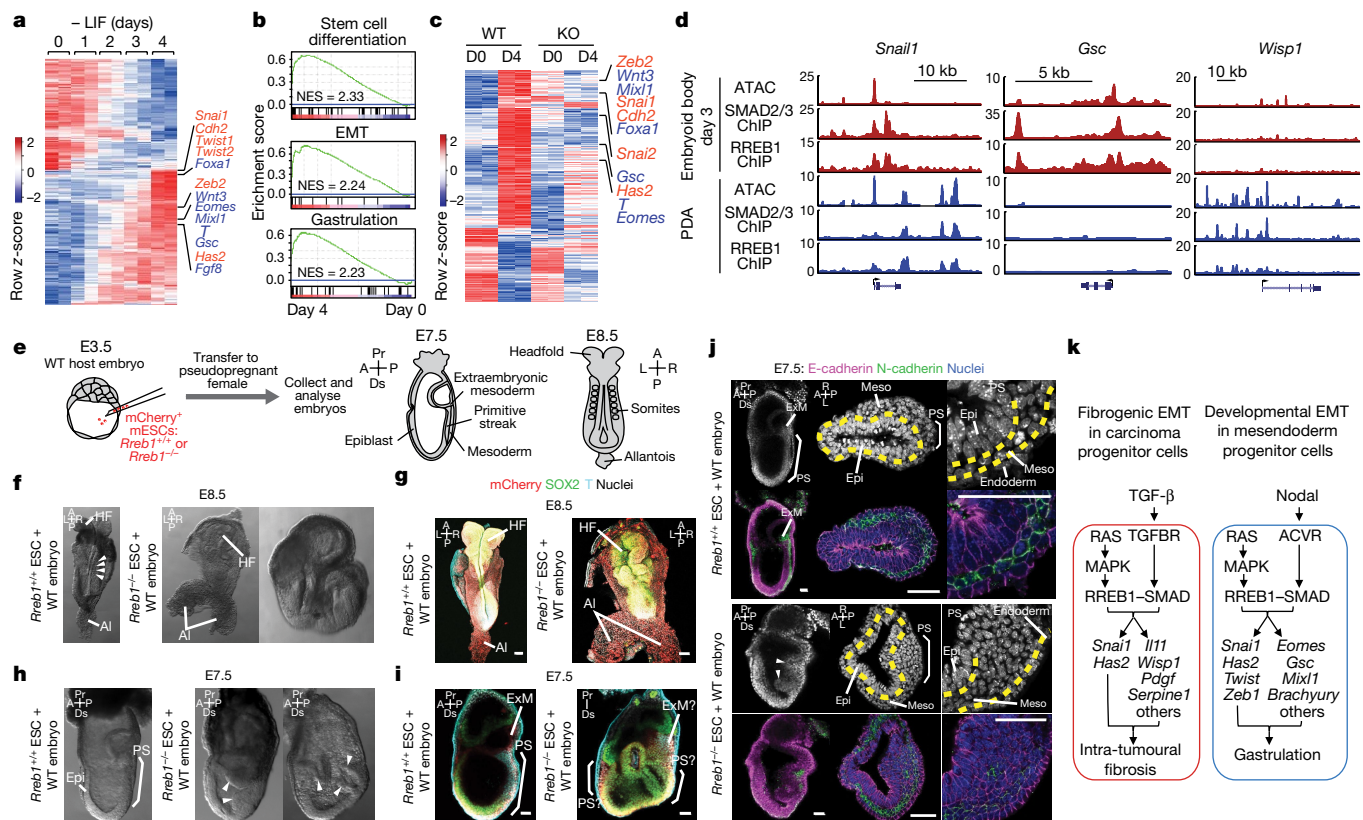
RNA-seq analysis of wild-type and *Rreb1*-knockout ES cells under pluripotency conditions (day 0) and after four days of embryoid body differentiation (day 4) showed few differences between wild-type and *Rreb1*-knockout cells on day 0, but lack of differentiation on day 4 (Fig. 4c), together with an absence of signatures of stem cell differentiation, EMT and gastrulation gene signatures (Extended Data Fig. 8f). Nodal and activin receptors signal through SMAD2/3<sup>9</sup>. SMAD2/3 ChIP-seq peaks in day 3 embryoid bodies overlapped with HA-RREB1 ChIP peaks genome-wide (Fig. 4d, Extended Data Fig. 9a–c), providing evidence for direct cooperation of SMADs and RREB1 in mesendoderm differentiation and EMT.

The assay for transposase-accessible chromatin using sequencing (ATAC-seq) revealed a shared major peak of chromatin accessibility on the *Snai1* promoter<sup>43</sup> in embryoid bodies and PDA cells, which overlapped with ChIP-seq SMAD2/3 and RREB1 binding profiles (Fig. 4d). The ATAC-seq profile overlapped with the ChIP-seq profiles on differentiation genes in day 3 embryoid body, and with *Wisp1* and *Serpine1* in PDA cells (Fig. 4d, Extended Data Fig. 9c). ATAC-seq revealed low chromatin accessibility at *Gsc* and *Mixl1* in PDA cells and at *Wisp1* and *Serpine1* in embryoid bodies, suggesting that different chromatin accessibility patterns enable SMAD2/3 and RREB1 access to *Snai1* and *Has2*, but with contextual restriction from fibrogenic and mesendoderm loci.

### RREB1 requirement during gastrulation

To determine whether RREB1 regulates gastrulation in vivo, we assessed the development of chimeric embryos comprising *Rreb1*<sup>-/-</sup> ES cells (Fig. 4e). Whereas *Rreb1*<sup>+/-</sup> chimaeras generally developed normally, the majority (approximately 75%) of *Rreb1*<sup>-/-</sup> ES-cell-containing embryos exhibited severe morphological abnormalities (Extended Data Fig. 10a, b). At E8.5, we observed aberrant development of neuroectoderm, comprising irregular neural plate folding (Fig. 4f, g) and disproportionate and bilaterally asymmetric headfolds (Extended Data Fig. 10c), defective intersomitic boundaries (Fig. 4f), and ectopic somite-like structures (Extended Data Fig. 10d). Some mutant chimaeras were so defective that specific structures, including the primitive streak and anterior–posterior axis, could not be discerned (Fig. 4f). We also noted axis duplications, including duplications of the epiblast (Extended Data Fig. 10c); posterior derivatives, including the allantois (Fig. 4f); and anterior derivatives, including the headfolds (Extended Data Fig. 10c, e).

At E7.5, approximately 75% of mutant embryo chimaeras were developmentally retarded or morphologically abnormal (Extended Data Fig. 10a–b, f). Similar to wild-type embryos, chimaeras containing wild-type ES cells formed a primitive streak and expressed markers of differentiation and EMT (Fig. 4h–j, Extended Data Fig. 10f). Although mutant embryo chimaeras expressed T and SNAIL within the primitive streak and nascent mesoderm (in both wild-type and *Rreb1*<sup>-/-</sup> cells), they frequently showed an accumulation of cells in the posterior epiblast, resulting in bulges into the amniotic cavity and/or a folded epiblast layer containing multiple cavities (Fig. 4h, i, Extended Data Fig. 10f, g), defects characteristic of gastrulation failure. No difference was detected in the number of mitotic or apoptotic cells between wild-type and mutant embryo chimaeras (Extended Data Fig. 10h, i).



**Fig. 4 | RREB1 and SMAD regulate distinct context-dependent EMTs.** **a**, Heat map of regulated transcripts during embryoid body differentiation. RNA-seq was performed at indicated times after shifting ES cells into differentiation medium (without leukaemia inhibitory factor (LIF)). EMT (red) and mesoderm lineage genes (blue) are highlighted. *n* = 2. **b**, Gene set enrichment analysis for the indicated signatures in day 4 embryoid bodies. NES, normalized enrichment score. **c**, Regulated transcripts (fold change >4 or <0.25) in WT and *Rreb1*-KO cells, on day 4 relative to day 0 of differentiation. *n* = 2. **d**, Gene track view of ATAC-seq, and SMAD2/3 and RREB1 ChIP-seq tags at indicated loci, in day 3 embryoid bodies (red tracks) versus TGF- $\beta$ -treated (1.5 h) PDA cells (blue tracks). ATAC-seq and ChIP-seq were performed once and confirmed for selected regions by qPCR of selected genomic regions. **e**, Chimaeras generated by injecting WT *Rreb1*<sup>+/+</sup> or mutant *Rreb1*<sup>-/-</sup> mCherry-tagged ES cells into WT mouse blastocysts were transferred to pseudopregnant females and dissected at E7.5–E8.5. **f, h**, Bright-field images of WT and *Rreb1*<sup>-/-</sup> chimeric embryos at E8.5 (**f**) and E7.5 (**h**). *Rreb1*<sup>-/-</sup> chimaeras displayed morphological defects. Arrowheads; somites (**f**), abnormal accumulation of

cells within epiblast (**h**). **g, i, j**, Confocal images of whole-mount immunostained chimaeras. **g**, Maximum intensity projections of E8.5 chimaeras showing abnormal neur ectoderm and axis duplication in *Rreb1*<sup>-/-</sup> chimaera. **i**, Sagittal section showing ectopic Brachyury expression, extensive epiblast folding and multiple cavities in *Rreb1*<sup>-/-</sup> chimaera (right). Anterior–posterior orientation of the embryo was not possible. **j**, Sagittal sections of whole chimaeras and representative sections through primitive streak region. Arrowheads, abnormal epiblast folding. Yellow dashed lines, epiblast–mesoderm boundary. Brackets, primitive streak. HF, headfold; NT, neural tube; AI, allantois; Epi, epiblast; PS, primitive streak; ExM, extraembryonic mesoderm; meso, mesoderm; A, anterior; P, posterior; Pr, proximal; Ds, distal; L, left; R, right. Scale bars, 50  $\mu$ m. Images in **f–j** are representative of two independent experiments. **k**, Summary of RAS-dependent TGF- $\beta$  or Nodal effects, coordinately triggered by cooperation between RREB1 and SMAD2/3 to activate EMT and associated contextual programs in carcinoma progenitors and pluripotent embryonic cells. Main target genes in each program and context are indicated. See also Extended Data Figs. 8–10.

In wild-type embryos and chimaeras, there was a switch of E-cadherin to N-cadherin as cells ingressed through the primitive streak (Fig. 4j). In mutant embryo chimaeras, cells within the aberrant bulges or folds continued to express E-cadherin and either did not strongly upregulate N-cadherin (Fig. 4j) or co-expressed both cadherins, with some embryos exhibiting ectopic N-cadherin within the posterior epiblast (Extended Data Fig. 10j). Together, these data demonstrate that mutant cells do not undergo a proper EMT at the primitive streak, resulting in gastrulation defects. Notably, *Rreb1*-knockout cells did not exhibit an absolute EMT block. Considering EMT as a continuum of states<sup>4</sup>, several EMT states—including Nodal- and RREB1-dependent EMTs—may overlap temporally and spatially within the embryo<sup>44,45</sup>.

## Discussion

The present work reveals how TGF- $\beta$  and RAS–MAPK signals acting jointly through SMAD and RREB1 transcription factors trigger EMTs in

different contexts (Fig. 4k). EMT and mesoderm differentiation are entwined events during gastrulation<sup>1,4,46</sup>, and our results shed light on how they are linked. SMADs, via RREB1, directly regulate the expression of EMT transcription factors and mesoderm genes in pluripotent progenitors, and of EMT transcription factors and fibrogenic factors in carcinoma cells. The induction of SNAIL and fibrogenic mediators are biologically coordinated but experimentally separable processes. This level of coordination is distinct from, and adds to the role of, SNAIL as inducer of downstream fibrogenic signals in renal fibrosis<sup>47,48</sup>. EMTs can couple to either morphogenic or fibrogenic events depending on context, and our evidence points at an epigenetic basis for this contextual nature of EMTs. With 15 zinc fingers and large interdomain regions, RREB1 probably coordinates interactions between DNA, SMAD proteins and other cofactors<sup>42,49,50</sup>. RREB1 is an understudied RAS effector, the structural and functional properties and genetic alterations of which warrant further attention. The generality of the TGF- $\beta$ –SMAD–RREB1 mechanism as a trigger of diverse EMTs provides

common ground for the analysis of EMTs in developmental and regenerative processes and paves the way for a better understanding of the role of TGF- $\beta$  in the pathogenesis of organ fibrosis and cancer.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-019-1897-5>.

- Arnold, S. J. & Robertson, E. J. Making a commitment: cell lineage allocation and axis patterning in the early mouse embryo. *Nat. Rev. Mol. Cell Biol.* **10**, 91–103 (2009).
- Dongre, A. & Weinberg, R. A. New insights into the mechanisms of epithelial–mesenchymal transition and implications for cancer. *Nat. Rev. Mol. Cell Biol.* **20**, 69–84 (2019).
- Ferretti, E. & Hadjantonakis, A. K. Mesoderm specification and diversification: from single cells to emergent tissues. *Curr. Opin. Cell Biol.* **61**, 110–116 (2019).
- Nieto, M. A., Huang, R. Y., Jackson, R. A. & Thiery, J. P. EMT: 2016. *Cell* **166**, 21–45 (2016).
- Battle, E. et al. The transcription factor snail is a repressor of E-cadherin gene expression in epithelial tumour cells. *Nat. Cell Biol.* **2**, 84–89 (2000).
- Cano, A. et al. The transcription factor snail controls epithelial–mesenchymal transitions by repressing E-cadherin expression. *Nat. Cell Biol.* **2**, 76–83 (2000).
- Brabletz, S. & Brabletz, T. The ZEB/miR-200 feedback loop—a motor of cellular plasticity in development and cancer? *EMBO Rep.* **11**, 670–677 (2010).
- De Craene, B. & Berr, G. Regulatory networks defining EMT during cancer initiation and progression. *Nat. Rev. Cancer* **13**, 97–110 (2013).
- David, C. J. & Massagué, J. Contextual determinants of TGF $\beta$  action in development, immunity and cancer. *Nat. Rev. Mol. Cell Biol.* **19**, 419–435 (2018).
- Heldin, C. H., Vanlandewijck, M. & Moustakas, A. Regulation of EMT by TGF $\beta$  in cancer. *FEBS Lett.* **586**, 1959–1970 (2012).
- Meng, X. M., Nikolic-Paterson, D. J. & Lan, H. Y. TGF- $\beta$ : the master regulator of fibrosis. *Nat. Rev. Nephrol.* **12**, 325–338 (2016).
- David, C. J. et al. TGF- $\beta$  tumor suppression through a lethal EMT. *Cell* **164**, 1015–1030 (2016).
- Horiguchi, K. et al. Role of Ras signaling in the induction of Snail by transforming growth factor- $\beta$ . *J. Biol. Chem.* **284**, 245–253 (2009).
- Janda, E. et al. Ras and TGF $\beta$  cooperatively regulate epithelial cell plasticity and metastasis: dissection of Ras signaling pathways. *J. Cell Biol.* **156**, 299–314 (2002).
- Meno, C. et al. Mouse Lefty2 and zebrafish antivin are feedback inhibitors of Nodal signaling during vertebrate gastrulation. *Mol. Cell* **4**, 287–298 (1999).
- Oft, M., Akhurst, R. J. & Balmain, A. Metastasis is driven by sequential elevation of H-ras and Smad2 levels. *Nat. Cell Biol.* **4**, 487–494 (2002).
- Sun, X., Meyers, E. N., Lewandoski, M. & Martin, G. R. Targeted disruption of *Fgf8* causes failure of cell migration in the gastrulating mouse embryo. *Genes Dev.* **13**, 1834–1846 (1999).
- Yamaguchi, T. P., Harpal, K., Henkemeyer, M. & Rossant, J. *fgfr-1* is required for embryonic growth and mesodermal patterning during mouse gastrulation. *Genes Dev.* **8**, 3032–3044 (1994).
- Zhou, X., Sasaki, H., Lowe, L., Hogan, B. L. & Kuehn, M. R. *Nodal* is a novel TGF- $\beta$ -like gene expressed in the mouse node during gastrulation. *Nature* **361**, 543–547 (1993).
- Cancer Genome Atlas Research Network. Integrated genomic characterization of pancreatic ductal adenocarcinoma. *Cancer Cell* **32**, 185–203 (2017).
- Thiagalingam, A. et al. RREB-1, a novel zinc finger protein, is involved in the differentiation response to Ras in human medullary thyroid carcinomas. *Mol. Cell. Biol.* **16**, 5335–5345 (1996).
- DaCosta Byfield, S., Major, C., Laping, N. J. & Roberts, A. B. SB-505124 is a selective inhibitor of transforming growth factor- $\beta$  type I receptors ALK4, ALK5, and ALK7. *Mol. Pharmacol.* **65**, 744–752 (2004).
- Porsch, H. et al. Efficient TGF $\beta$ -induced epithelial–mesenchymal transition depends on hyaluronan synthase HAS2. *Oncogene* **32**, 4355–4365 (2013).
- Martin-Malpartida, P. et al. Structural basis for genome wide recognition of 5-bp GC motifs by SMAD transcription factors. *Nat. Commun.* **8**, 2070 (2017).
- Costello, L. C., Zou, J., Desouki, M. M. & Franklin, R. B. Evidence for changes in RREB-1, ZIP3, and Zinc in the early development of pancreatic adenocarcinoma. *J. Gastrointest. Cancer* **43**, 570–578 (2012).
- Kent, O. A., Fox-Talbot, K. & Halushka, M. K. RREB1 repressed miR-143/145 modulates KRAS signaling through downregulation of multiple targets. *Oncogene* **32**, 2576–2585 (2013).
- Yamane, T. et al. Transcriptional activation of the cholecystokinin gene by DJ-1 through interaction of DJ-1 with RREB1 and the effect of DJ-1 on the cholecystokinin level in mice. *PLoS ONE* **8**, e78374 (2013).
- Kent, O. A. et al. Repression of the miR-143/145 cluster by oncogenic Ras initiates a tumor-promoting feed-forward pathway. *Genes Dev.* **24**, 2754–2759 (2010).
- Winslow, M. M. et al. Suppression of lung adenocarcinoma progression by Nkx2-1. *Nature* **473**, 101–104 (2011).
- Kasai, H., Allen, J. T., Mason, R. M., Kamimura, T. & Zhang, Z. TGF- $\beta$ 1 induces human alveolar epithelial to mesenchymal cell transition (EMT). *Respir. Res.* **6**, 56 (2005).
- Schafer, S. et al. IL-11 is a crucial determinant of cardiovascular fibrosis. *Nature* **552**, 110–115 (2017).
- Toda, N., Mukoyama, M., Yanagita, M. & Yokoi, H. CTGF in kidney fibrosis and glomerulonephritis. *Inflamm. Regen.* **38**, 14 (2018).
- Königshoff, M. et al. WNT1-inducible signaling protein-1 mediates pulmonary fibrosis in mice and is upregulated in humans with idiopathic pulmonary fibrosis. *J. Clin. Invest.* **119**, 772–781 (2009).
- Yoshida, S. et al. Extrahepatic platelet-derived growth factor- $\beta$ , delivered by platelets, promotes activation of hepatic stellate cells and biliary fibrosis in mice. *Gastroenterology* **147**, 1378–1392 (2014).
- Dave, N. et al. Functional cooperation between Snail1 and twist in the regulation of ZEB1 expression during epithelial to mesenchymal transition. *J. Biol. Chem.* **286**, 12024–12032 (2011).
- Ye, X. et al. Distinct EMT programs control normal mammary stem cells and tumour-initiating cells. *Nature* **525**, 256–260 (2015).
- Scheel, C. et al. Paracrine and autocrine signals induce and maintain mesenchymal and stem cell states in the breast. *Cell* **145**, 926–940 (2011).
- Miettinen, P. J., Ebner, R., Lopez, A. R. & Derynck, R. TGF- $\beta$  induced transdifferentiation of mammary epithelial cells to mesenchymal cells: involvement of type I receptors. *J. Cell Biol.* **127**, 2021–2036 (1994).
- Shirakihara, T., Saitoh, M. & Miyazono, K. Differential regulation of epithelial and mesenchymal markers by  $\delta$ EF1 proteins in epithelial mesenchymal transition induced by TGF- $\beta$ . *Mol. Biol. Cell* **18**, 3533–3544 (2007).
- Xie, L. et al. Activation of the Erk pathway is required for TGF- $\beta$ 1-induced EMT *in vitro*. *Neoplasia* **6**, 603–610 (2004).
- Peng, G. et al. Spatial transcriptome for the molecular annotation of lineage fates and cell identity in mid-gastrula mouse embryo. *Dev. Cell* **36**, 681–697 (2016).
- Wang, Q. et al. The p53 family coordinates Wnt and Nodal inputs in mesendodermal differentiation of embryonic stem cells. *Cell Stem Cell* **20**, 70–86 (2017).
- Peiró, S. et al. Snail1 transcriptional repressor binds to its own promoter and controls its expression. *Nucleic Acids Res.* **34**, 2077–2084 (2006).
- Lee, J. D., Silva-Gagliardi, N. F., Tepass, U., McGlade, C. J. & Anderson, K. V. The FERM protein Epb4.115 is required for organization of the neural plate and for the epithelial–mesenchymal transition at the primitive streak of the mouse embryo. *Development* **134**, 2007–2016 (2007).
- Zohn, I. E. et al. p38 and a p38-interacting protein are critical for downregulation of E-cadherin during mouse gastrulation. *Cell* **125**, 957–969 (2006).
- Ciruna, B. & Rossant, J. FGF signaling regulates mesoderm cell fate specification and morphogenetic movement at the primitive streak. *Dev. Cell* **1**, 37–49 (2001).
- Grande, M. T. et al. Snail1-induced partial epithelial-to-mesenchymal transition drives renal fibrosis in mice and can be targeted to reverse established disease. *Nat. Med.* **21**, 989–997 (2015).
- Lovisa, S. et al. Epithelial-to-mesenchymal transition induces cell cycle arrest and parenchymal damage in renal fibrosis. *Nat. Med.* **21**, 998–1009 (2015).
- Yoon, S. J., Wills, A. E., Chuong, E., Gupta, R. & Baker, J. C. HEB and E2A function as SMAD/FOXH1 cofactors. *Genes Dev.* **25**, 1654–1661 (2011).
- Thuault, S. et al. HMGA2 and Smads co-regulate SNAIL1 expression during induction of epithelial-to-mesenchymal transition. *J. Biol. Chem.* **283**, 33437–33446 (2008).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2020

## Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

## Data availability

Source Data for Fig. 2d, i, j and Extended Data Figs. 2a, d, 3b, c, e, 4c, d, 5c, i, 6d, k, 7c are provided with the online version of the paper. The ChIP-seq and RNA-seq data have been deposited in the Gene Expression Omnibus under accession numbers GSE118765 and GSE128958. All other data are available from the authors on reasonable request.

**Acknowledgements** We thank L. Tian and D.-F. Lee for technical assistance, L. Huang for pLenti-HA-Rreb1 plasmid and K. Anderson, A. Nieto and J. P. Thiery for insightful discussions. We acknowledge the support of Y. Furuta and S. Gong of the Mouse Genetics Core, the Molecular Cytology, Integrated Genomics and Flow Cytometry Cores of MSKCC, and S. Y. Kim of the Rodent Genetic Engineering Laboratory of NYU. This work was supported by NIH grants R01CA34610 (J.M.), P01-CA129243 (J.M.), R01DK084391 (A.-K.H.), R01HD094868 (A.-K.H.) and

P30-CA008748 (MSKCC). J.S. was supported by an AACR Basic Cancer Research Fellowship (16-40-01-SUJ1) and a Charles H. Revson Senior Fellowship in Biomedical Science (17-23). S.M.M. was supported by a Sir Henry Wellcome Postdoctoral Fellowship. Y.-H.H. was supported by a Medical Scientist Training Program grant (T32GM007739) and Predoctoral Fellowship (F30-CA203238) from the National Cancer Institute. H.B. was supported by a Damon Runyon Postdoctoral Fellowship.

**Author Contributions** J.S. and J.M. conceived the project and designed the study. J.M., J.S. and A.-K.H. wrote the manuscript. J.S. performed most experiments. S.M.M. and A.-K.H. performed embryo analyses. C.J.D. performed shRNA screens. Q.W. generated E14 *Rreb1*-knockout clones. E.E.E. and H.B. performed molecular cloning of RREB1 phosphorylation-mutant constructs. Y.Z. and Y.-H.H. performed ATAC-seq. W.S. assisted with tissue culture, molecular biology and mouse experiments. J.S., R.C.H. and R.K.S. performed stable isotope labelling with amino acids in cell culture (SILAC) experiments. J.S. and Y.-H.H. performed bioinformatics analyses.

**Competing interests** J.M. serves in the scientific advisory board and owns company stock in Scholar Rock.

## Additional information

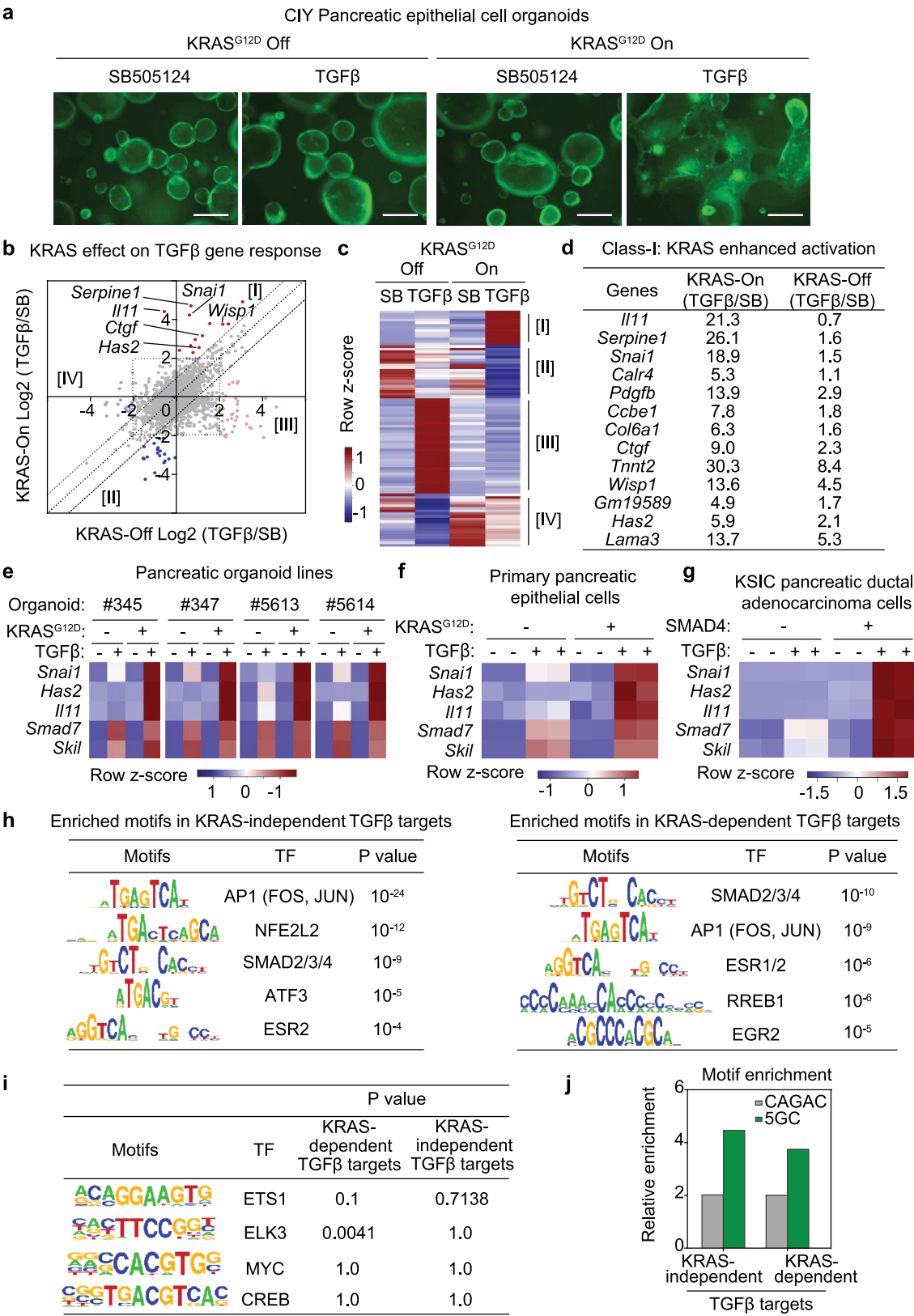
**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41586-019-1897-5>.

**Correspondence and requests for materials** should be addressed to J.M.

**Peer review information** Nature thanks Aristidis Moustakas, Angela Nieto and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>.



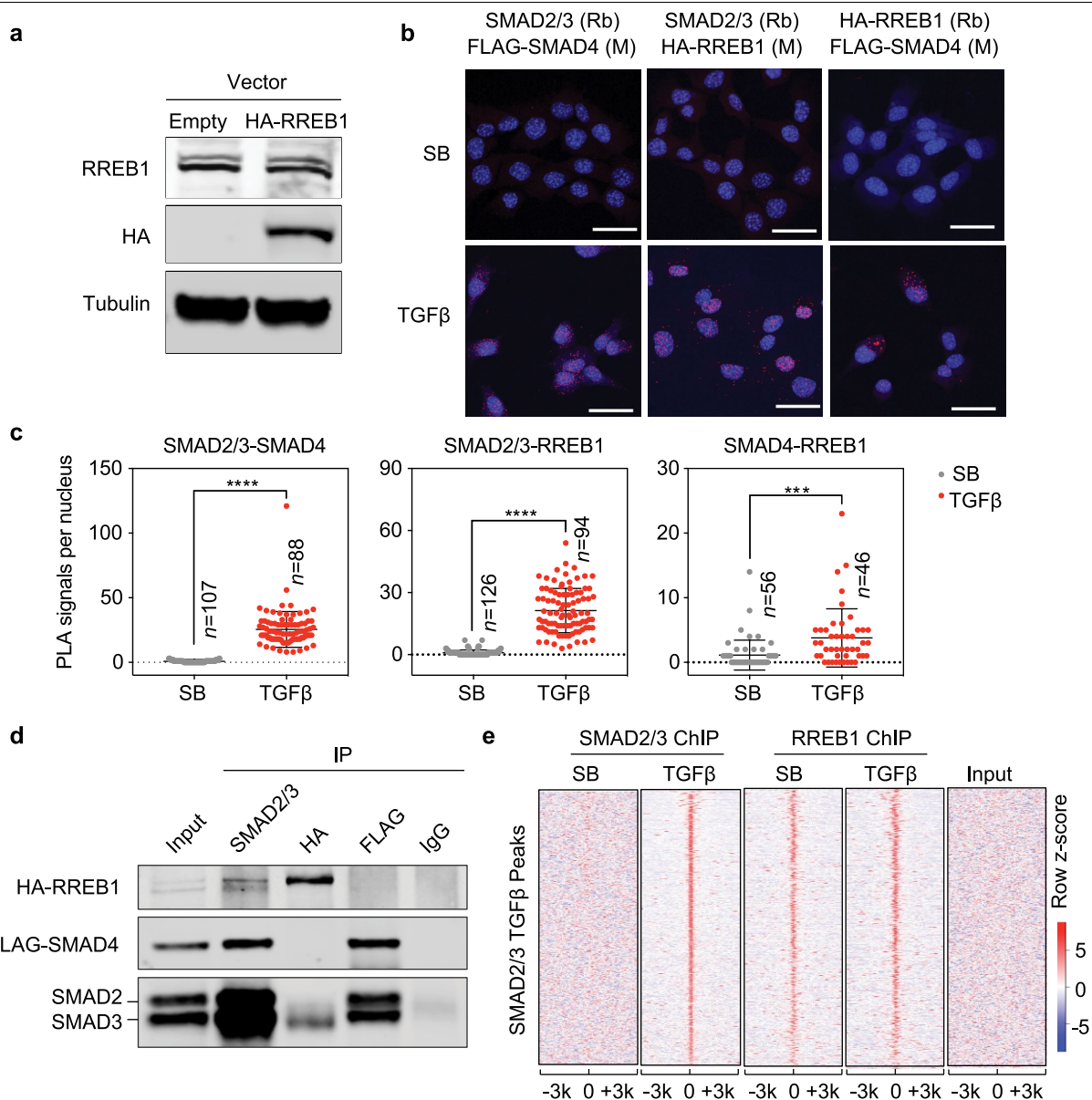




## Extended Data Fig. 1 | RREB1 as a SMAD cofactor in TGF- $\beta$ gene responses.

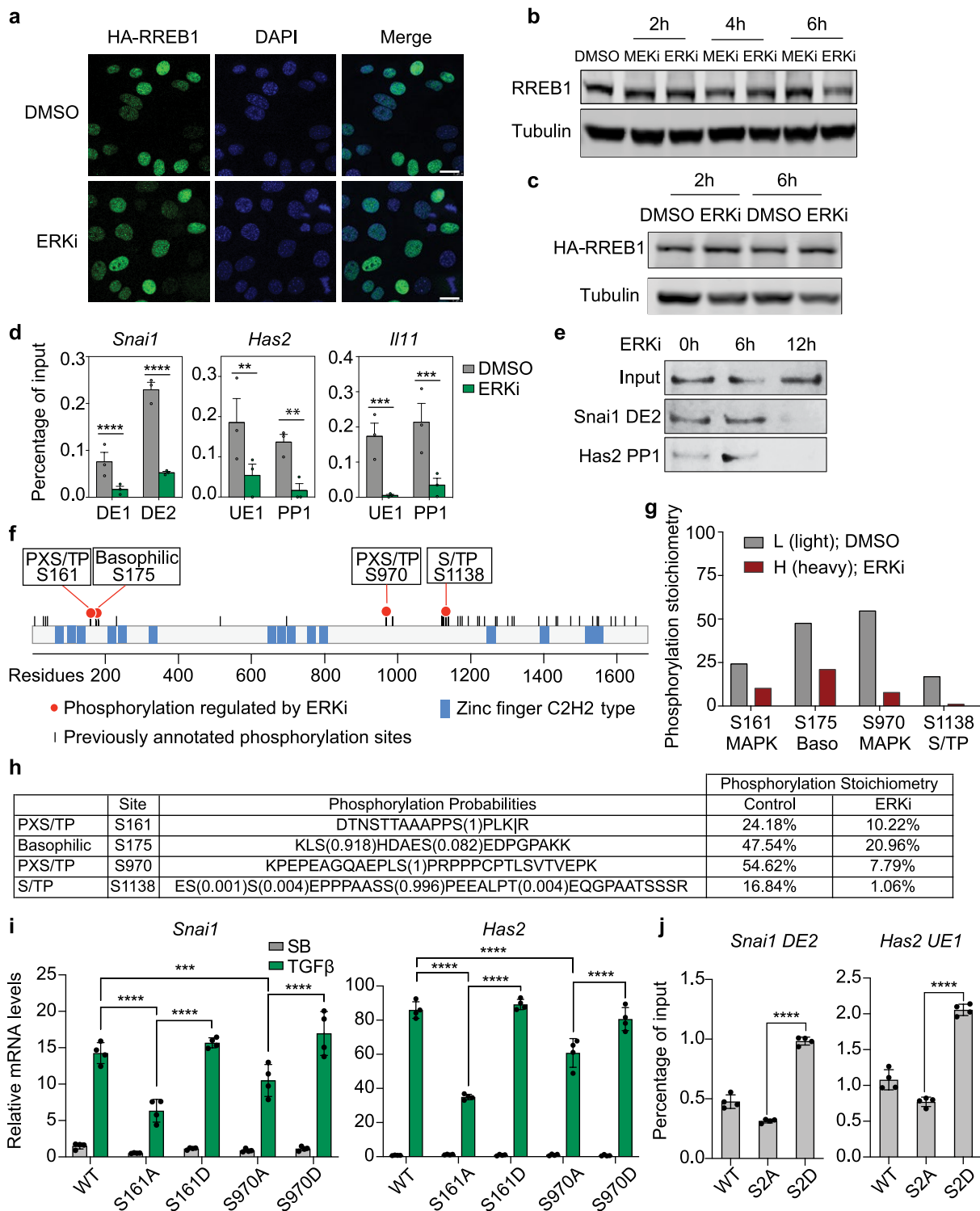
**a**, YFP fluorescence images of CIY organoids expressing KRAS(G12D) under doxycycline control treated with SB505124 or TGF- $\beta$  for 2.5 days. Scale bars, 200  $\mu$ m. Images are representative of two independent experiments. **b**, Influence of KRAS(G12D) on TGF- $\beta$  gene responses. CIY pancreatic organoids inducibly expressing KRAS(G12D), were treated with SB505124 or TGF- $\beta$  for 1.5 h and analysed by RNA-seq. Dots represent fold change (in  $\log_2$ ) in mRNA levels of individual genes under TGF- $\beta$  versus SB505124 treatment conditions, with KRAS(G12D) expression turned off (x axis) or on (y axis). Off-diagonal dots correspond to TGF- $\beta$  gene responses that were enabled (groups I and II) or disabled (groups III and IV) by KRAS(G12D). Gene activation (I and III) and repression responses (II and IV) are included. **c**, Heat map of four classes of KRAS-modified TGF- $\beta$  gene responses.  $n = 1$ . Representative result of two independent experiments. Classes I–IV correspond to the off-diagonal genes derived from the RNA-seq in **b**. **d**, TGF- $\beta$  gene activation responses augmented by KRAS(G12D) (class I responses) in CIY pancreatic organoids. Fold increase in mRNA levels in TGF- $\beta$  versus SB505124 treatment conditions in presence or absence of inducible KRAS(G12D). **e**, Heat maps showing TGF- $\beta$  induction of *Snai1*, *Has2*, *Il11*, *Smad7* and *Skil* in four independent CIY mouse pancreatic organoid lines with inducible KRAS(G12D) expression.  $n = 4$ . **f**, Heat map of the

indicated TGF- $\beta$  gene responses in spheroid cultures of pancreatic epithelial cells (PECs) inducibly expressing KRAS(G12D).  $n = 2$ . **g**, Heat map of the indicated TGF- $\beta$  gene responses in monolayer cultures of mouse *Kras*<sup>G12D</sup>, *Smad4*<sup>fl/jf1</sup>, *Cdkn2a*<sup>fl/jf1</sup>, *Pdx1-cre* (KSIC) PDA cell lines transduced with a SMAD4 vector or an empty vector.  $n = 2$ . **h**, Transcription factor (TF)-binding motifs enriched in KRAS-independent SMAD2/3 binding sites (left) and KRAS-dependent SMAD2/3 binding sites (right). SMAD2/3 ChIP-seq analyses were performed in SMAD4-restored PDA cells that were treated with SB505124 (2.5  $\mu$ M) or TGF- $\beta$  (100 pM) for 1.5 h. Transcription factor binding-motif analyses were performed with PscanChIP.  $n = 821$  peak regions (left).  $n = 778$  peak regions (right). **i**, Motif enrichment analysis of RAS-regulated transcription factors in KRAS-dependent ( $n = 778$  peak regions) and KRAS-independent ( $n = 821$  peak regions) SMAD2/3 binding sites. **j**, Comparative enrichment of classic SMAD binding motifs (CAGAC and GGCTG) and 5GC motifs (GGC(GC)|(CG)) in a 200-bp region of SMAD2/3 ChIP peaks within 1,000 bp of a transcriptional start site<sup>24</sup>. The relative enrichment is normalized to the baseline dataset obtained from 20,000 random 200-bp regions from the mm10 genome assembly. The 5GC motifs are enriched approximately fourfold in SMAD2/3 ChIP peaks compared to the baseline, and the classic motifs are enriched twofold.



**Extended Data Fig. 2 | RREB1 interacts with SMAD and binds to TGF- $\beta$  target genes. **a**, Western blot analysis of RREB1 and HA-RREB1 levels in SMAD4-restored PDA cells stably transduced with a HA-RREB1 vector. Tubulin immunoblotting was used as loading control. Data are representative of two independent experiments. **b**, Proximity ligation assay showing TGF- $\beta$ -dependent proximity between RREB1, SMAD2/3 and SMAD4 in the nucleus. Scale bars, 30  $\mu$ m. Data are representative of two independent experiments. **c**, Quantification of PLA signals in **b**. Cell numbers ( $n$ ) of each group are indicated in the graph, two-tailed unpaired  $t$ -test. Data are mean  $\pm$  s.d. \*\*\*\* $P$  < 0.0001, \*\*\* $P$  < 0.001. **d**, SMAD4-restored PDA cells expressing HA-RREB1**

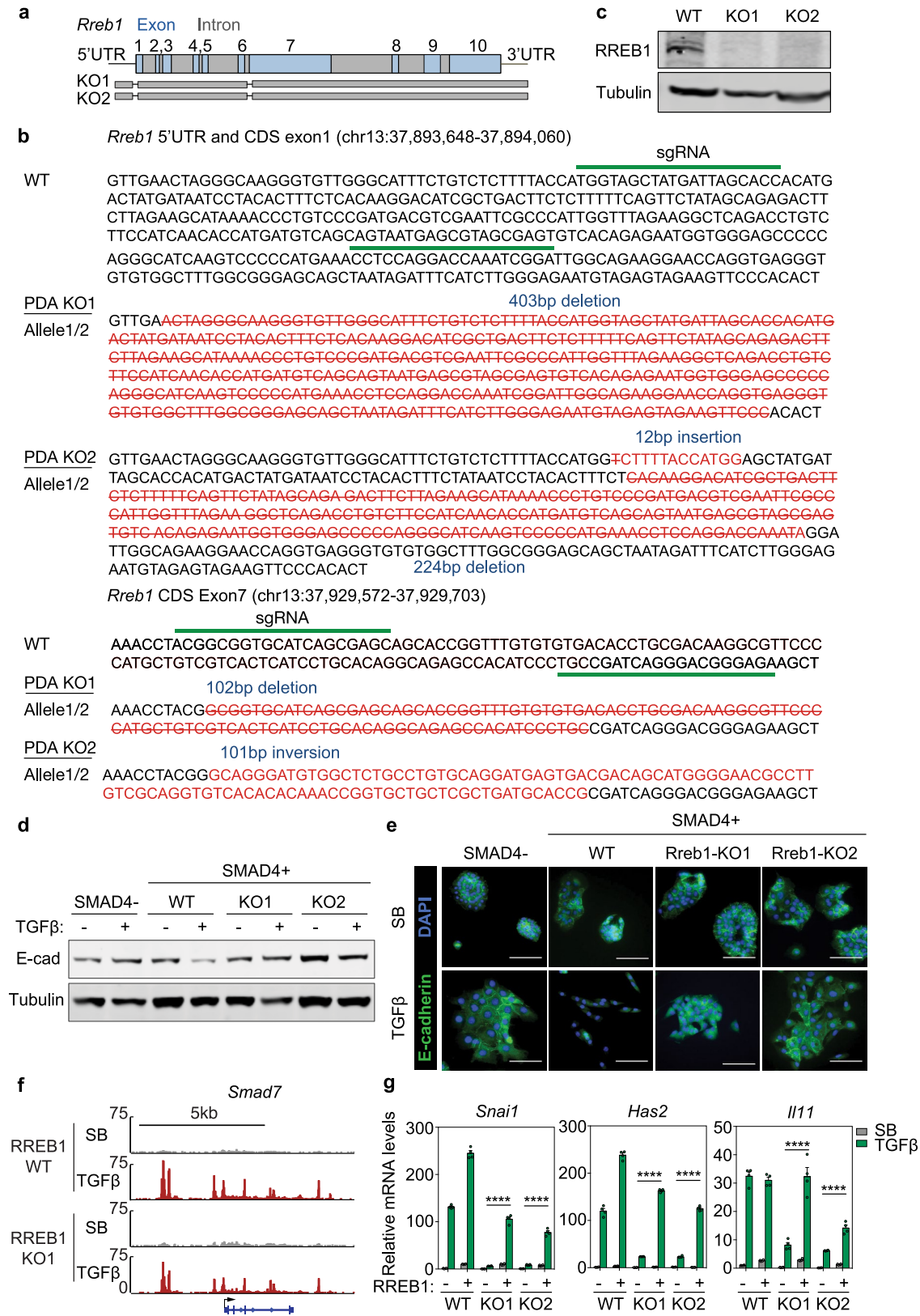
were treated with TGF- $\beta$  for 1.5 h, lysed and immunoprecipitated (IP) with the indicated antibodies. The immune complexes were collected and subjected to western blot with the antibodies indicated on the left. Data are representative of two independent experiments. **e**, Heat map of ChIP-seq tag densities for SMAD2/3 and HA-RREB1 in genomic regions  $\pm$ 3 kb from the centre of SMAD2/3 binding peaks in SMAD4-restored PDA cells that were treated with SB505124 or TGF- $\beta$  for 1.5 h and subjected to SMAD2/3 and HA-RREB1 ChIP-seq analysis. ChIP-seq was performed once, and an independent ChIP was performed in which selective genomic regions were confirmed by qPCR.



Extended Data Fig. 3 | See next page for caption.

**Extended Data Fig. 3 | RREB1 is phosphorylated and regulated by ERK. a,** Representative immunofluorescence images of HA-RREB1 in SMAD4-restored PDA cells treated with DMSO or 1  $\mu$ M ERK inhibitor SCH772984 (ERKi) for 6 h. Scale bar, 20  $\mu$ m. Data are representative of two independent experiments. **b, c,** Western blot analysis of RREB1 (**b**) or HA-RREB1 levels (**c**) in SMAD4-restored PDA cells treated with DMSO, 1  $\mu$ M ERKi or 1  $\mu$ M AZD6244 (MEKi; an inhibitor of the ERK-activating kinases MEK1/2) for the indicated time periods. Tubulin immunoblotting was used as loading control. Data are representative of two independent experiments. **d,** ChIP-PCR analysis of HA-RREB1 binding to the indicated sites (Figs. 1g, 3b) in *Snail*, *Has2* and *Il11* in SMAD4-restored PDA cells that were treated with vehicle (DMSO) or ERKi (1  $\mu$ M) for 6 h. Mean  $\pm$  s.e.m.  $n = 3$ , two-way ANOVA.  $**P < 0.01$ ,  $***P < 0.001$ ,  $****P < 0.0001$ . **e,** SMAD4-restored PDA cells expressing HA-RREB1 were treated with ERKi for the indicated length of time. HA-RREB1 was tested for binding to *Snail* DE2 and *Has2* PP1 double-stranded DNA oligonucleotide probes in DNA affinity precipitation assays. Data are representative of two independent experiments. **f,** Schematic of RREB1. Each tick represents a previously annotated

phosphorylation site in PhosphoSitePlus identified in at least two independent mass spectrometry experiments. Red filled circles represent high stoichiometry (>15%) phosphorylation sites that are inhibited by ERKi, as identified in **g**. Zinc-finger domains annotated in Uniprot are shown. **g,** Phosphorylation stoichiometry of four ERK-dependent RREB1 phosphorylation sites in SMAD4-restored PDA cells, as determined by SILAC mass spectrometry of cells treated with DMSO (control) in light medium or ERKi in heavy medium for 6 h. **h,** Summary of ERK-dependent RREB1 phosphorylation sites, sequence motifs and phosphorylation stoichiometry. **i,** *Rreb1*-KO PDA cells were transduced with the indicated RREB1-WT or phosphorylation-site mutant constructs, then treated with SB505124 or TGF- $\beta$  for 1.5 h. mRNA levels of *Snail* and *Has2* were determined by qPCR with reverse transcription. Mean  $\pm$  s.e.m.  $n = 4$ , two-way ANOVA.  $***P < 0.001$ ,  $****P < 0.0001$ . **j,** ChIP-PCR analysis of HA-RREB1 binding to the indicated sites in *Rreb1*-KO PDA cells transduced with the indicated RREB1-WT or phosphorylation-site mutant constructs. Mean  $\pm$  s.e.m.  $n = 4$ , two-tailed unpaired *t*-test.  $****P < 0.0001$ .

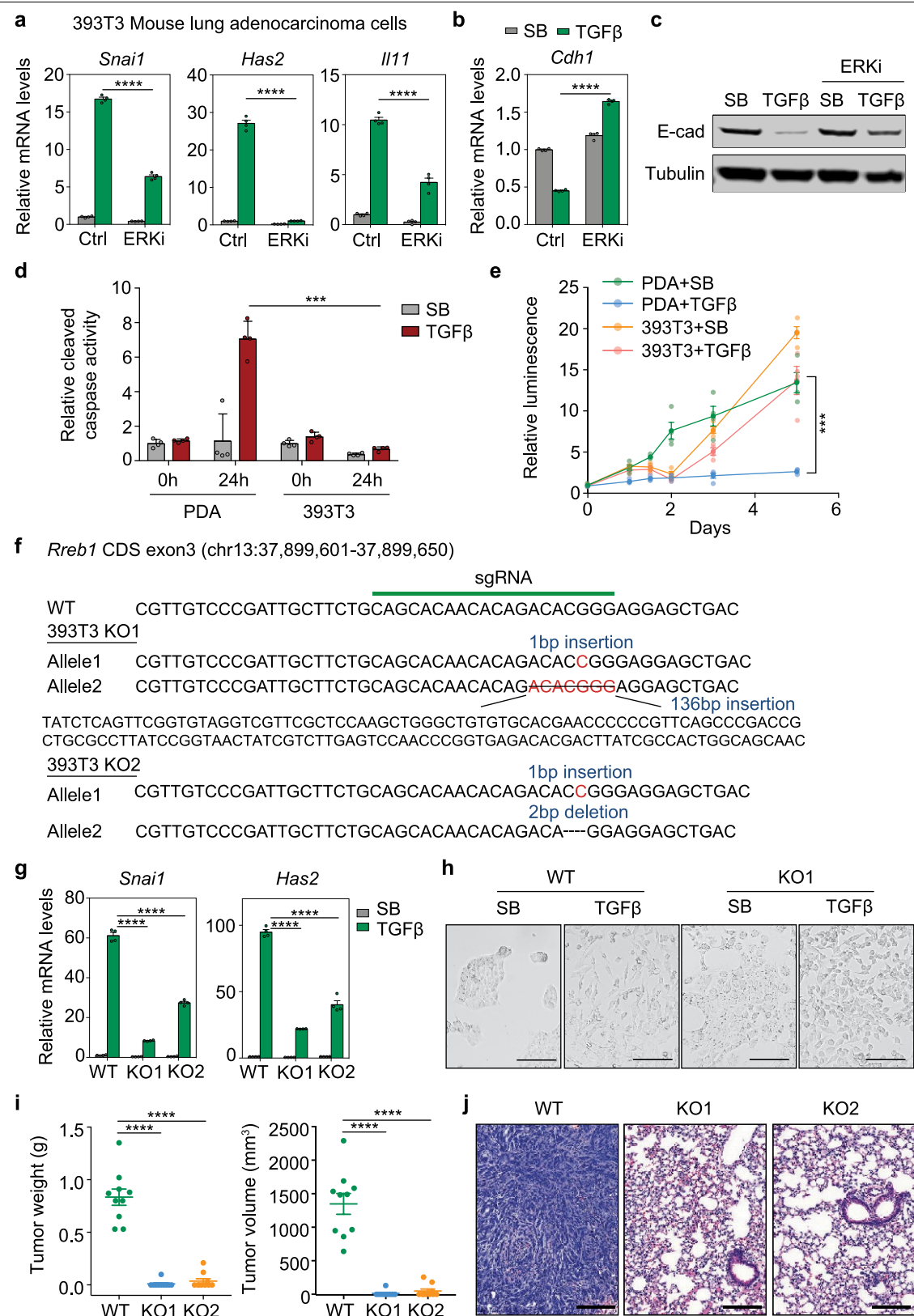


Extended Data Fig. 4 | See next page for caption.



**Extended Data Fig. 4 | RREB1 mediates KRAS-dependent TGF- $\beta$  responses in PDA cells.** **a**, Scheme of CRISPR–Cas9-mediated mutation of *Rreb1* in mouse SMAD4-restored PDA cells. **b**, sgRNA sequences and genomic sequences of *Rreb1* coding region (CDS) exons 1 and 7 in mutant clones KO1 and KO2 derived from SMAD4-restored PDA cells. **c**, Western blot analysis of RREB1 levels in WT and *Rreb1*-KO cells. Tubulin immunoblotting was used as loading control. Data are representative of two independent experiments. **d**, Western blot analysis of E-cadherin in mouse KSIC PDA cells, SMAD4-restored PDA cells and two *Rreb1*-KO SMAD4-restored PDA clones, treated with SB505124 or TGF- $\beta$  for 24 h. Tubulin immunoblotting was used as loading control. Data are representative

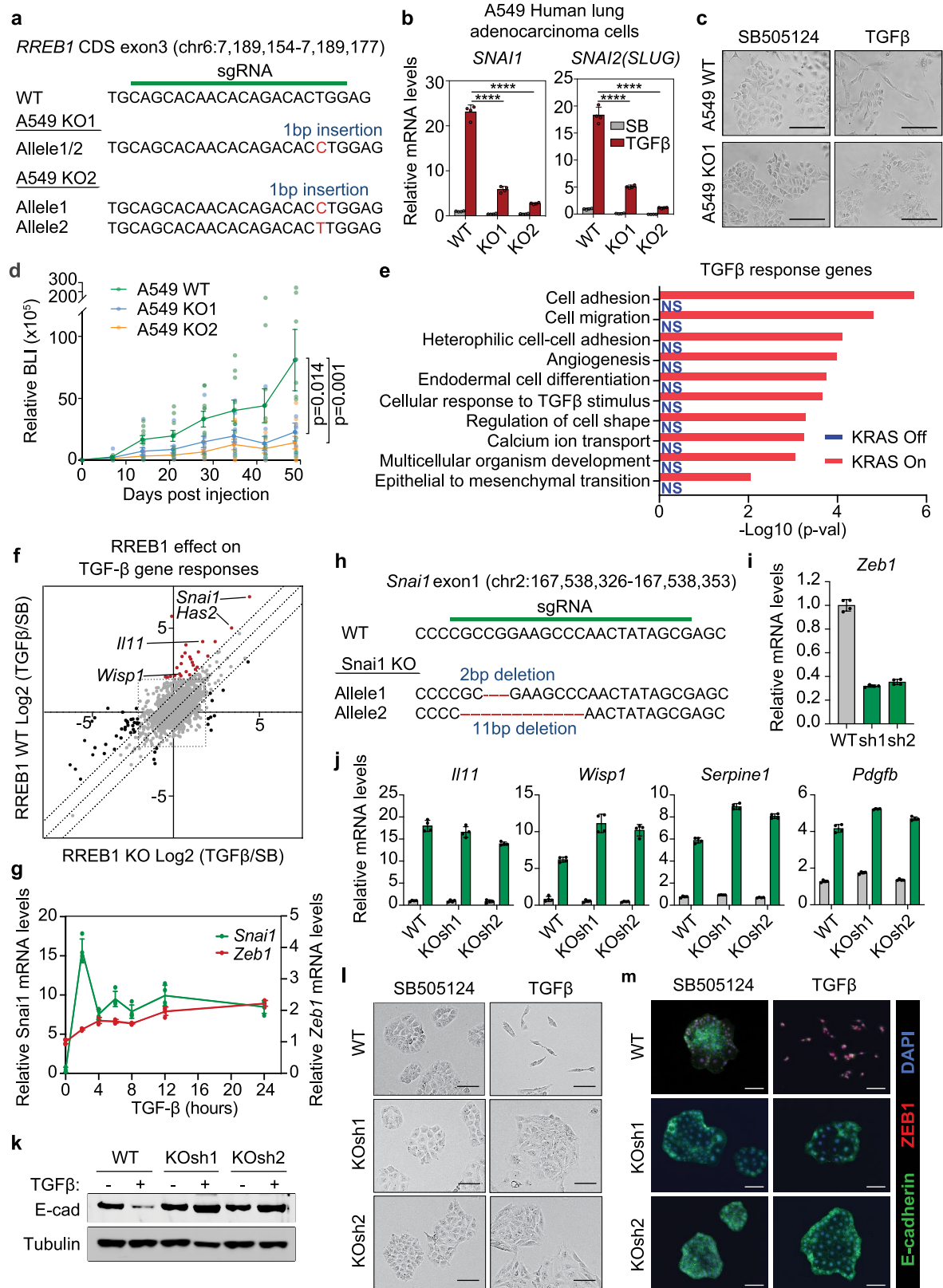
of two independent experiments. **e**, Representative E-cadherin immunofluorescence and DAPI staining of the same cells as in **d** treated with SB505124 or TGF- $\beta$  for 48 h. Scale bars, 100  $\mu$ m. Data are representative of two independent experiments. **f**, Gene track view of SMAD2/3 ChIP-seq tags in the *Smad7* locus of the WT and *Rreb1*-KO PDA cells. The gene body is schematically represented at the bottom. ChIP-seq was performed once and an independent ChIP was performed in which selective genomic regions were confirmed by qPCR. **g**, mRNA levels of *Snai1*, *Has2* and *Il11* in WT and two *Rreb1*-KO cells that were transduced with an RREB1 vector or empty vector and then treated with SB505124 or TGF- $\beta$  for 1.5 h. Mean  $\pm$  s.e.m.  $n = 4$ ; two-way ANOVA; \*\*\*\* $P < 0.0001$ .



Extended Data Fig. 5 | See next page for caption.

**Extended Data Fig. 5 | RREB1 mediates tumorigenic EMT in lung adenocarcinoma cells.** **a**, *Snai1*, *Has2* and *Il11* mRNA levels in 393T3 mouse LUAD cells treated with DMSO (Ctrl) or ERKi (SCH772984, 1  $\mu$ M) for 6 h, followed by treatment of SB505124 or TGF- $\beta$  for 1.5 h. Mean  $\pm$  s.e.m.  $n = 4$ ; two-way ANOVA; \*\*\*\* $P < 0.0001$ . **b**, *Cdh1* mRNA levels in 393T3 cells with the indicated treatments for 48 h. Mean  $\pm$  s.d.  $n = 4$ ; two-way ANOVA. **c**, Western blot analysis of E-cadherin in 393T3 cells with the indicated treatments for 48 h. Tubulin immunoblotting was used as loading control. Data are representative of two independent experiments. **d**, SMAD4-restored PDA cells and 393T3 LUAD cells cultured in D10F containing 2.5  $\mu$ M MK2206<sup>12</sup> were treated with SB505124 (2.5  $\mu$ M) or TGF- $\beta$  (100 pM) and assayed for cleaved caspase 3/7 activity at the indicated times. Mean  $\pm$  s.e.m.  $n = 4$ ; two-way ANOVA; \*\*\*\* $P < 0.001$ . **e**, SMAD4-restored PDA cells and 393T3 cells cultured in D10F

containing 2.5  $\mu$ M MK2206 were treated with SB505124 or TGF- $\beta$ . Cell viability was determined at the indicated times. Mean  $\pm$  s.e.m.  $n = 4$ ; two-way ANOVA; \*\*\*\* $P < 0.001$ . **f**, sgRNA sequence targeting *Rreb1* CDS exon 3, and mutant *Rreb1* genomic sequences of the resulting 393T3 KO1 and KO2 clones. **g**, mRNA levels of *Snai1* and *Has2* in the WT and *Rreb1*-KO 393T3 cells after treatment with SB505124 (2.5  $\mu$ M) or TGF- $\beta$  (100 pM) for 1.5 h. Mean  $\pm$  s.e.m.  $n = 4$ ; two-way ANOVA; \*\*\*\* $P < 0.0001$ . **h**, Phase contrast images of 393T3 cell monolayers treated with SB505124 or TGF- $\beta$  for 48 h. Scale bars, 200  $\mu$ m. Data are representative of two independent experiments. **i**, Weight and volume of tumours in Fig. 2i. Mean  $\pm$  s.e.m.  $n = 10$ , two sites were inoculated per mouse; two-tailed unpaired  $t$ -test; \*\*\*\* $P < 0.0001$ . **j**, Representative haematoxylin and eosin staining images of indicated lung tissue sections from Fig. 2j. Scale bars, 200  $\mu$ m. Data are representative of two independent experiments.

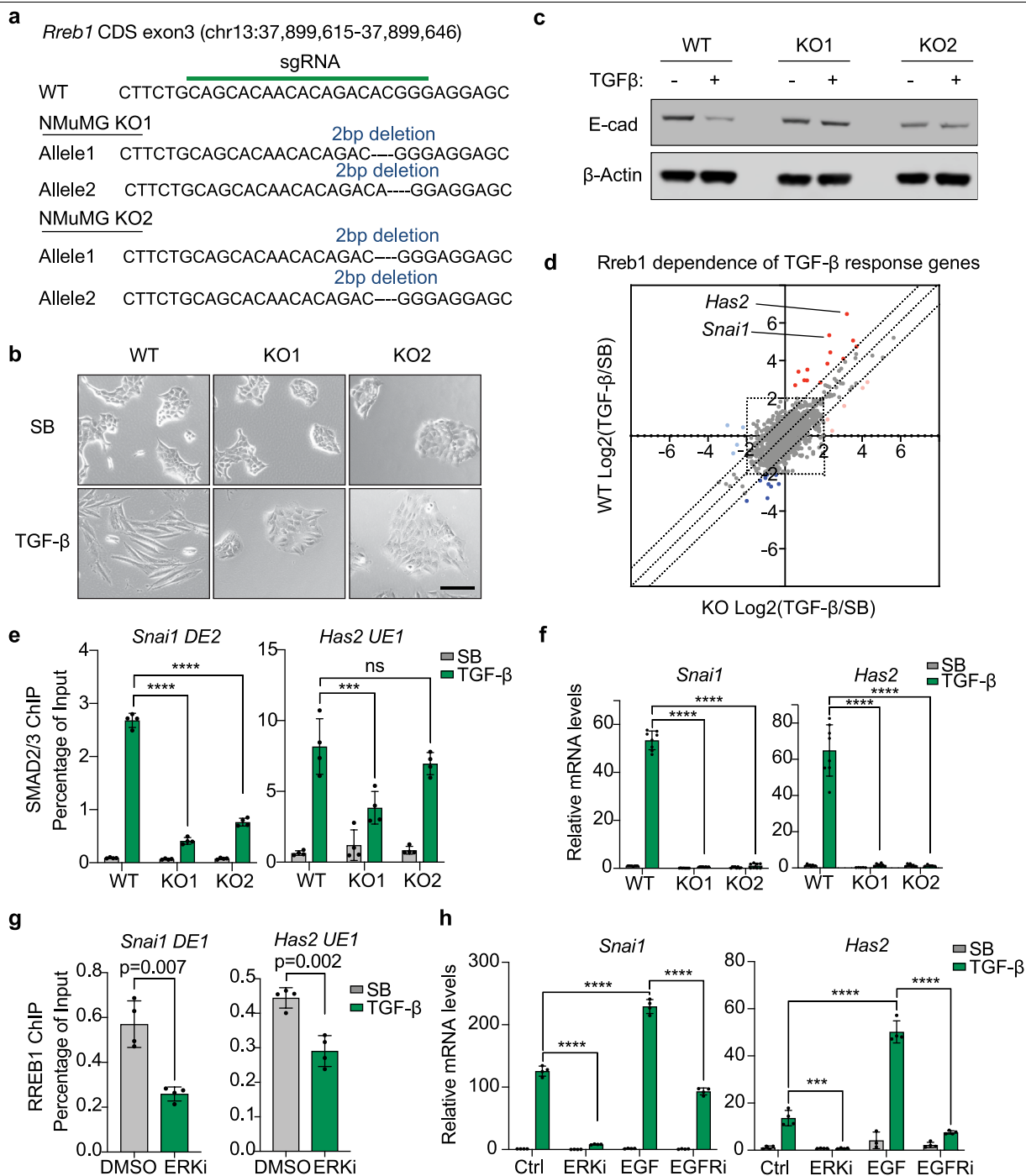


Extended Data Fig. 6 | See next page for caption.

**Extended Data Fig. 6 | RREB1-dependent TGF- $\beta$  responses in LUAD and PDA cells.** **a**, sgRNA sequence targeting *RREB1* CDS exon 3 and mutant *RREB1* genomic sequences of the resulting A549 KO1 and KO2 clones. **b**, *SNAIL* and *SLUG* mRNA levels in WT A549 and two *RREB1* KO clones treated with SB505124 or TGF- $\beta$  for 24 h. Mean  $\pm$  s.e.m.  $n = 4$ ; two-way ANOVA; \*\*\*\* $P < 0.0001$ . **c**, Phase-contrast images of WT A549 and *RREB1*-KO cell monolayers treated with SB505124 or TGF- $\beta$  for 48 h. Scale bars, 200  $\mu$ m. Data are representative of two independent experiments. **d**, Growth kinetics of tumours formed by subcutaneously inoculated WT or *RREB1*-KO A549 cells in athymic mice, as determined by BLI of a transduced firefly luciferase gene in the cells. Mean  $\pm$  s.e.m.  $n = 10$ , two sites were inoculated per mouse; two-way ANOVA. **e**, Gene ontology analysis of TGF- $\beta$  response genes in CIY organoids inducibly expressing KRAS(G12D), based on the RNA-seq in Extended Data Fig. 1b. **f**, WT and *Rreb1*-KO PDA cells were treated with SB505124 or TGF- $\beta$  for 1.5 h and

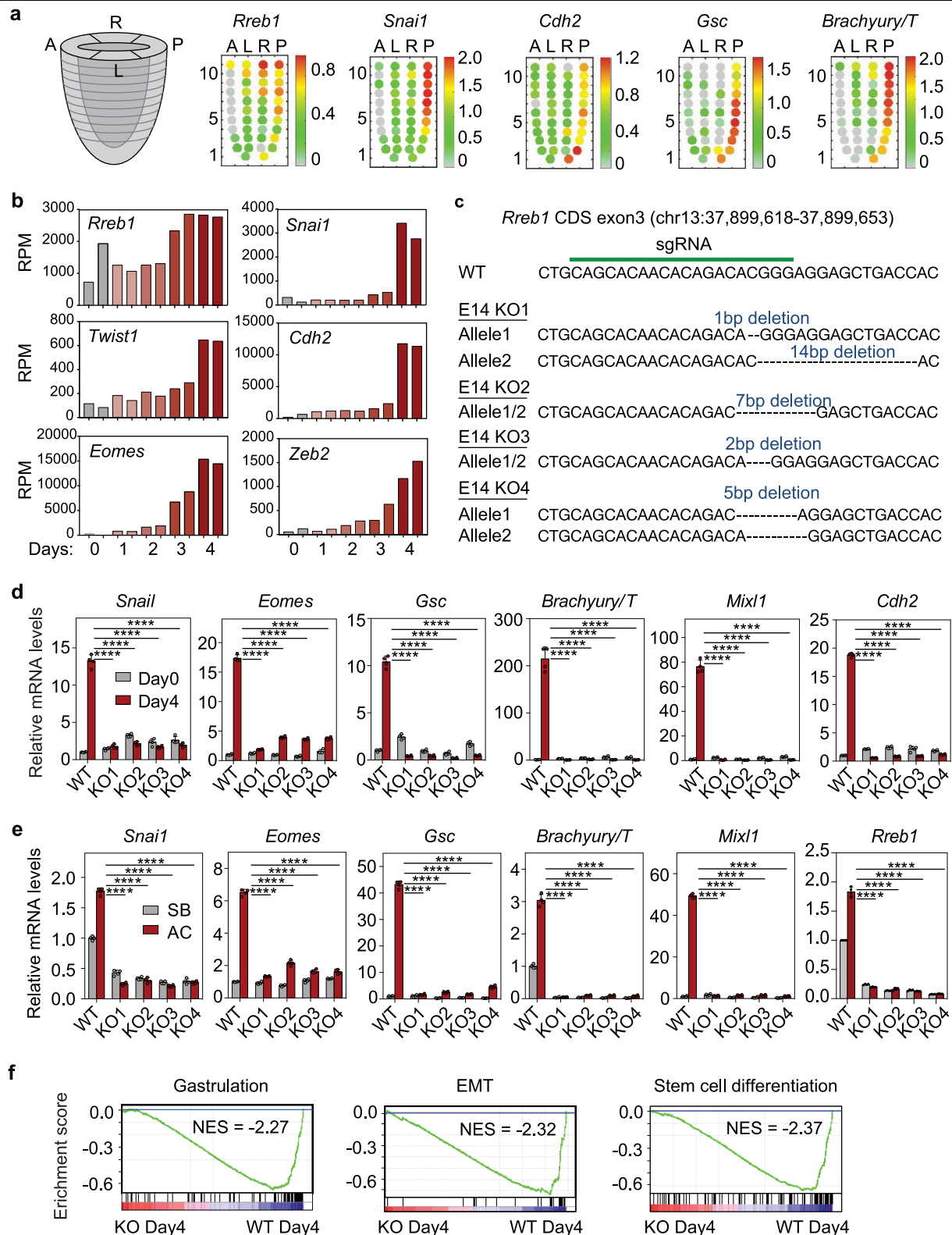
analysed by RNA-seq. Dots represent fold change (in log<sub>2</sub>) in mRNA levels of individual genes under TGF- $\beta$  versus SB505124 treatment conditions, in *Rreb1*-KO (x axis) or WT cells (y axis). Off-diagonal dots corresponding to *Snail*, *Has2*, *Illi* and *Wisp1* are highlighted. **g**, Induction of *Snail* and *Zeb1* expression by TGF- $\beta$  in mouse PDA cells. Mean  $\pm$  s.d.  $n = 4$ . **h**, sgRNA sequence targeting *Snail* and resulting mutant *Snail* genomic sequences in mouse PDA cells. **i**, Knockdown of *Zeb1* with two independent shRNAs in SNAIL-KO mouse PDA cells (KOsh cells). Mean  $\pm$  s.d.  $n = 4$ . **j**, Fibrogenic gene responses to TGF- $\beta$  in WT and SNAIL and ZEB1-double depleted KOsh PDA cells. Mean  $\pm$  s.d.  $n = 4$ . **k-m**, E-cadherin levels (**k**), phase-contrast images (**l**) and E-cadherin and Zeb1 immunofluorescence in WT and KOsh PDA cells that were treated with SB505124 or TGF- $\beta$  for 48 h. Scale bars, 100  $\mu$ m. Data are representative of two independent experiments.





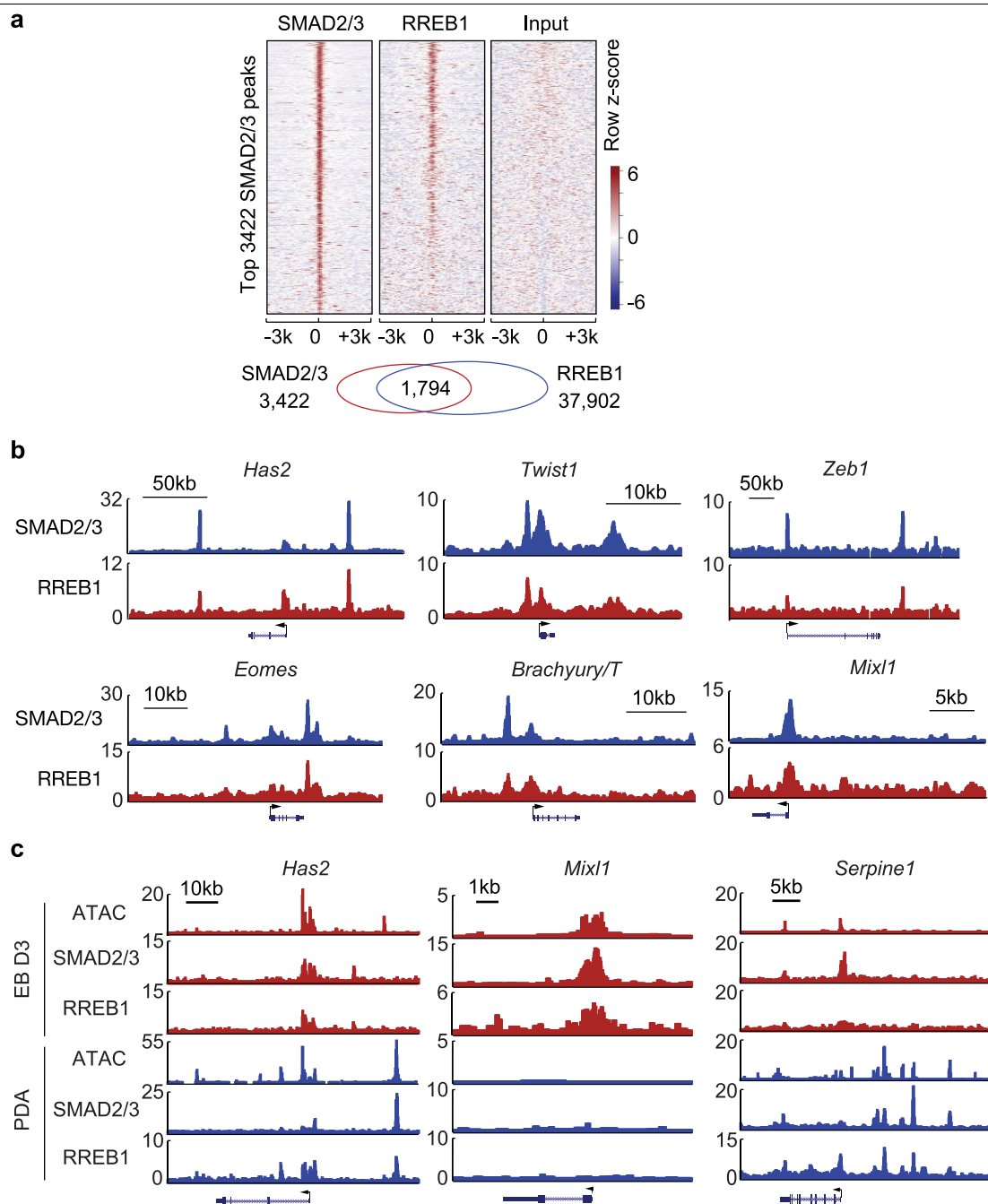
**Extended Data Fig. 7 | RREB1-dependent TGF- $\beta$  responses in mammary epithelial cells. a**, sgRNA sequence targeting *RREB1* CDS exon 3 and mutant *RREB1* genomic sequences of the resulting NMuMG KO1 and KO2 clones. **b**, Phase-contrast images of WT and *Rreb1*-KO NMuMG cell monolayers treated with SB505124 or TGF- $\beta$  for 48 h. Scale bar, 100  $\mu$ m. Data are representative of two independent experiments. **c**, Western blot analysis of E-cadherin in WT and *Rreb1*-KO NMuMG cells, treated with SB505124 or TGF- $\beta$  for 48 h.  $\beta$ -actin immunoblotting was used as loading control. Data are representative of two independent experiments. **d**, WT and *Rreb1*-KO NMuMG cells were treated with SB505124 or TGF- $\beta$  for 1.5 h and analysed by RNA-seq. Dots represent fold change (in log<sub>2</sub>) in mRNA levels of individual genes under TGF- $\beta$  versus SB505124 treatment conditions, in *Rreb1*-KO (x axis) or WT cells (y axis). Off-diagonal dots corresponding to *Snai1* and *Has2* are highlighted. **e**, ChIP-PCR

analysis of SMAD2/3 binding to the *Snai1* (DE2) and *Has2* (UE1) regions (Fig. 1g) in WT and *Rreb1*-KO NMuMG cells. Cells were treated with 2.5  $\mu$ M SB505125 or 100 pM TGF- $\beta$  for 1.5 h. Mean  $\pm$  s.e.m.  $n$  = 4; two-way ANOVA; \*\*\* $P$  < 0.001, \*\*\*\* $P$  < 0.0001. **f**, mRNA levels of *Snai1* and *Has2* in WT and *Rreb1*-KO NMuMG cells after treatment with SB505124 or TGF- $\beta$  for 1.5 h. Mean  $\pm$  s.e.m.  $n$  = 4; two-way ANOVA; \*\*\*\* $P$  < 0.0001. **g**, ChIP-PCR analysis of HA-RREB1 binding to the indicated *Snai1* and *Has2* regions in NMuMG cells that were treated with vehicle (DMSO) or the ERKi SCH772984 (1  $\mu$ M) for 6 h. Mean  $\pm$  s.e.m.  $n$  = 3; two-tailed unpaired  $t$ -test. **h**, *Snai1* and *Has2* mRNA levels in NMuMG cells treated with DMSO (Ctrl), ERKi (1  $\mu$ M SCH772984), EGF (10 ng ml<sup>-1</sup>, 10 min) or EGFR inhibitor (gefitinib, 1  $\mu$ M, 2 h), followed by SB505124 or TGF- $\beta$  treatment for another 1.5 h. Mean  $\pm$  s.e.m.  $n$  = 4; two-way ANOVA; \*\*\* $P$  < 0.001, \*\*\*\* $P$  < 0.0001.



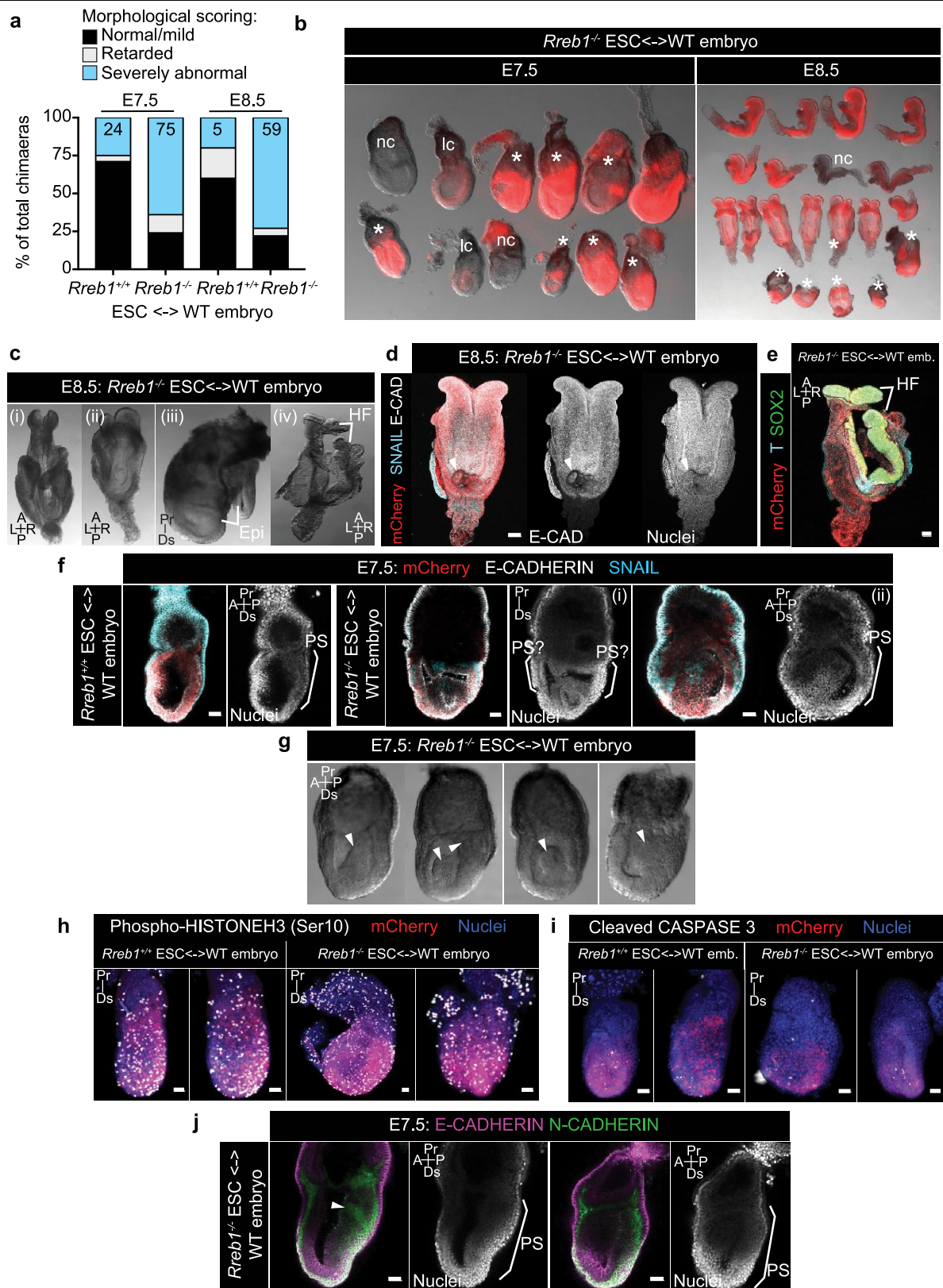
**Extended Data Fig. 8 | RREB1 in gastrulation EMT and mesendoderm differentiation.** **a**, Corn plot presentation of *Rreb1*, *Snai1*, *Cdh2*, *Gsc* and *T* in E7.0 mouse embryo. A, anterior; L, left; R, right; P, posterior regions. Each dot represents transcript level at a specific positional address. Heat map denotes expression level of each gene computed from transcript counts in RNA-seq datasets<sup>41</sup>. **b**, Reads per million reads (RPM) of *Rreb1*, *Snai1*, *Twist1*, *Cdh2*, *Eomes* and *Zeb2* in the RNA-seq dataset at the indicated times after shifting ES cells into LIF-deficient embryoid body differentiation medium. **c**, sgRNA sequence targeting *Rreb1* CDS exon 3, and mutant *Rreb1* genomic sequences of four

resulting mouse ES cell KO clones. **d**, mRNA levels of EMT (*Snai1* and *Cdh2*) and mesendoderm differentiation genes (*Eomes*, *Gsc*, *T* and *Mixl1*) in WT and four independent *Rreb1*-KO clones on day 4 embryoid body differentiation. Mean  $\pm$  s.d.  $n = 4$ ; two-way ANOVA; \*\*\*\* $P < 0.0001$ . **e**, mRNA levels of the indicated genes in WT and four independent *Rreb1*-KO clones treated with receptor inhibitor (SB505125) or activin A (AC) for 2 h. Mean  $\pm$  s.e.m.  $n = 4$ ; two-way ANOVA; \*\*\*\* $P < 0.0001$ . **f**, Gene set enrichment analysis for gastrulation, EMT and stem cell differentiation genes in WT cells, and absence in *Rreb1*-KO cells at day 4 embryoid body differentiation.



**Extended Data Fig. 9 | RREB1 and SMAD contextually regulate EMT genes. a,** Heat map of ChIP-seq tag densities for SMAD2/3 and HA-RREB1 in genomic regions  $\pm 3$  kb from centre of 3,422 high-confidence SMAD2/3 binding peaks in day 3 embryoid bodies subjected to SMAD2/3 and HA ChIP-seq analyses. **b,** Gene track view of SMAD2/3 and HA-RREB1 ChIP-seq tags in the loci of EMT genes (*Has2*, *Twist1* and *Zeb1*) and early mesendoderm lineage genes (*Eomes*, *T*

and *Mixl1*) in day 3 embryoid bodies. Gene bodies are schematically represented at the bottom of each track set. **c,** Gene track view of ATAC-seq and SMAD2/3 and RREB1 ChIP-seq tags on indicated loci, in day 3 embryoid bodies (red tracks) versus TGF- $\beta$  treated (1.5 h) SMAD4-restored PDA cells (blue tracks). In **a–c**, ChIP-seq was performed once and an independent ChIP was performed in which selective genomic regions were confirmed by qPCR.



**Extended Data Fig. 10** | See next page for caption.

# Article

**Extended Data Fig. 10 | *Rreb1*<sup>-/-</sup> mouse embryo chimaeras exhibit defects in early development.** **a**, E7.5 and E8.5 chimeric embryos containing WT ES cells or *Rreb1*<sup>-/-</sup> ES cells were scored, on the basis of gross morphology, as normal or mild defects, developmentally retarded or severely abnormal. At E7.5, a fraction of *Rreb1*<sup>+/-</sup> ES cell embryos displayed small clumps of cells in the amniotic cavity, possibly an artefact from the microinjection, and were therefore scored as abnormal. *Rreb1*<sup>-/-</sup> data are compiled from four distinct KO clones. **b**, Bright-field morphology and mCherry fluorescence (marking descendants of injected ES cells) in representative litters of *Rreb1*<sup>-/-</sup> ES-cell-containing chimeric embryos dissected at E7.5 and E8.5. nc, non-chimeric; lc, low chimaerism. Asterisks mark morphologically abnormal or developmentally retarded embryos. **c**, Bright-field images of morphologically abnormal *Rreb1*<sup>-/-</sup> ES-cell-containing chimeric E8.5 embryos. Embryos exhibited abnormal headfold development, including disproportionate headfolds (i) and asymmetric headfolds (ii). Axis duplication was also observed, (iii) and (iv). Of note, the embryo in (iii) is also developmentally retarded. **d, e**, Confocal maximum intensity projections of whole-mount

immunostained E8.5 *Rreb1*<sup>-/-</sup> ES-cell-containing chimeric embryos. **d**, An embryo with an ectopic somite-like structure (arrowhead). **e**, The embryo in **c** (iv) with axis duplication of the headfolds. **f**, Sagittal confocal optical sections of whole-mount immunostained chimeric E7.5 embryos. Embryos shown in **f** (i, ii) have multiple cavities and multiple expression sites of SNAIL, hence anterior–posterior axis orientation is not possible. **g**, Bright-field images of morphologically abnormal *Rreb1*<sup>-/-</sup> ES-cell-containing chimeric E7.5 embryos. Embryos frequently had protrusions into the cavity and thickening of the posterior epiblast, marked by arrowheads. **h, i**, Confocal maximum intensity projections of chimeric embryos after whole-mount immunostaining for phospho-histone H3 (**h**), labelling mitotic cells, and cleaved caspase 3 (**i**), labelling apoptotic cells. Brackets demarcate the primitive streak. **j**, Sagittal confocal optical sections of chimeric E7.5 embryos after whole-mount immunostaining for E-cadherin and N-cadherin. Arrowhead, aberrant N-cadherin expression. Scale bars, 50 µm. Images in **b–j** are representative of two independent experiments.



## Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

### Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a Confirmed

- ☐ ☒ The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- ☐ ☒ An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- ☐ ☒ The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- ☒ ☐ A description of all covariates tested
- ☐ ☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- ☐ ☒ A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- ☐ ☒ For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give  $P$  values as exact values whenever suitable.*
- ☒ ☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- ☒ ☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- ☒ ☐ Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated
- ☐ ☒ Clearly defined error bars  
*State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on [statistics for biologists](#) may be useful.

### Software and code

Policy information about [availability of computer code](#)

#### Data collection

BLI data was acquired using IVIS Spectrum Xenogen instrument (Caliper Life Sciences) using Living Image software v.2.50. Histology data for H&E and IHC were acquired using Mirax scanner. qRT-PCR data was acquired using ViiA 7 real-time PCR system, Applied Biosystems. Western blot data was acquired using Licor Image Studio V2. RNA-seq, ChIP-seq and ATAC-seq data have been deposited on GEO under the accession numbers provided in the methods section.

#### Data analysis

Statistical analysis: GraphPad Prism v 8.1.2  
Image processing and analysis: ImageJ v 2.0.0  
BLI data analysis: Living Image software v 2.50  
Western blot images and analysis: ImageStudioLite v 5.2.5  
RNA-seq, ChIP-seq and ATAC-seq analysis: FastQC v 0.11.5, GNUparallel v 2.5.2b, STAR v 0.6.1p1, HTSeq v 3.4, DESeq2 v 3.4, Bowtie2 v 2.3, Samtools v 0.1, R v 3.5.0, GSEA v 4.0.0, DAVID v 6.8, MACS v 1.4.2, HOMER v 4.10, Pscan-ChIP v 1.3. Custom code is available on request.  
Mass spectrometric data analysis: MaxQuant v 1.5.3.30.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

## Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The raw and processed RNA-seq, ChIP-seq and ATAC-seq data are deposited on GEO (GSE118765 and GSE128958):

<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE118765>

<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE128958>

The following secure token has been created to allow review of record:

GSE118765: ersqismjhypxex

GSE128958: ynupyawfdefdc

## Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/authors/policies/ReportingSummary-flat.pdf](https://www.nature.com/authors/policies/ReportingSummary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	For high throughput sequencing experiments, two independent biological replicates were used where possible and to strengthen the conclusions of the study. Sample sized for other assays (qPCR, ChIP-PCR, etc.) were chosen based on our prior experience and common standards in the field for detecting statistically significant differences between conditions.
Data exclusions	No data was excluded from the studies.
Replication	All attempts at replication were successful. Biological replicates of each experiment is stated under each figure legend and all attempts were successful. Moreover, findings were repeatedly reproduced throughout the study: RNA-seq with qRT-PCR in multiple cell lines; protein levels with Western blot and immunostaining; cell response assays with different cell line models.
Randomization	In animal experiments, mice were randomly assigned to each experiment group.
Blinding	Investigators were blinded to cell groups and mouse groups when collecting results.

## Reporting for specific materials, systems and methods

### Materials & experimental systems

n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Unique biological materials
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input type="checkbox"/>	<input checked="" type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants

### Methods

n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

## Unique biological materials

Policy information about [availability of materials](#)

Obtaining unique materials DNA plasmids: pLVX-Tight-Puro-KrasG12D, pLVX-IRES-Hyg-SMAD4 (Addgene #107128), SGEP (Addgene #111170), pLenti-HA-Rreb1, pSpCas9(BB)-2A-GFP (PX458) (Addgene #48138), pSpCas9(BB)-2A-Puro (PX459) (Addgene #48139), pLVX-EF1α-IRES-

## Antibodies

### Antibodies used

Antibody name Vendor Catalog # Lot # Dilution  
 Rabbit monoclonal anti-E-cadherin Cell Signaling Technology Cat #3195 Lot #13 1:1000  
 Rat monoclonal anti-E-cadherin Abcam Cat #ab11512 Lot #GR183496-1 1:1000  
 Rat monoclonal anti-E-cadherin Sigma Cat #U-3254 Lot #077M4800V 1:1000  
 Rabbit polyclonal anti-ZEB1 Santa cruz Cat #sc-25388 Lot #unknown 1:1000  
 Rabbit monoclonal anti-SMAD2/3 Cell Signaling Technology Cat #8685 Lot #6 1:1000  
 Rabbit monoclonal anti-HA-Tag Cell Signaling Technology Cat #3724 Lot #unknown 1:1000  
 Mouse monoclonal anti-FLAG M2 Sigma Cat #F1804 Lot #SLBT7654 1:2000  
 Rabbit polyclonal anti-RREB1 Genway Biotech Cat #GWB-5B0668 Lot #27022 1:500  
 Mouse monoclonal anti- $\alpha$ -Tubulin Sigma Cat #T6199 Lot #unkown 1:10,000  
 Rabbit polyclonal anti-Cleaved Caspase-3 Cell Signaling Technology Cat #9661 Lot #43 1:1000  
 Goat polyclonal anti-Brachyury R&D Cat #AF2085 Lot #unkown 1:200  
 Rabbit polyclonal anti-phospho-H3 Millipore Cat #06-570 Lot #6570 1:300  
 Rabbit polyclonal anti-N-cadherin Santa cruz Cat #sc-7939 Lot #unkown 1:300  
 Rabbit polyclonal anti-RFP Rockland Cat #600-401-379 Lot #42393 1:300  
 Goat polyclonal anti-SNAIL R&D Cat #AF3639 Lot #XRS0218061 1:100  
 Rat monoclonal anti-SOX2 eBioscience Cat #14-9811-82 Lot #2023691 1:200  
 Rabbit polyclonal anti-alpha SMA Abcam Cat #ab5694 Lot #GR248336-4 1:1000  
 Rabbit monoclonal [Y92] anti-PDGFR beta Abcam Cat #ab32570 Lot #GR212663-49 1:1000  
 Mouse monoclonal anti- $\beta$ -Actin Cell Signaling Technology Cat #3700 Lot #unknown 1:1000  
 Donkey anti-rabbit alexa 568 Invitrogen Cat #A10042 Lot #1964370 1:500  
 Donkey anti-goat alexa 488 Invitrogen Cat #A11055 Lot #830720 1:500  
 Donkey anti-rat dylight 650 Invitrogen Cat #SA5-10029 Lot #UF2789721 1:500  
 Donkey anti-rat alexa 488 Invitrogen Cat #A21208 Lot #1932496 1:500  
 Donkey anti-rabbit alexa 647 Invitrogen Cat #A31573 Lot #1903516 1:500  
 Donkey anti-rabbit alexa 488 Invitrogen Cat #A21206 Lot #2045215 1:500

### Validation

Validation statement for each primary antibody is provided on the manufacture's website.

## Eukaryotic cell lines

Policy information about [cell lines](#)

### Cell line source(s)

Mouse KSIC PDA cell line was provided by Nabeel Bardeesy. Mouse CIY pancreatic organoid lines were generated from CIY mice. Mouse lung epithelial cell line 393T3 cell line was a gift from Taylor Jacks. Other cell lines (human lung epithelial cell line A549, mouse mammary gland epithelial cell line NMuMG and mouse embryonic stem cell line E14Tg2a.IV) were purchased from ATCC.

### Authentication

Mouse KSIC PDA cell line, CIY pancreatic organoid lines and 393T3 cell line were genotyped by PCR amplification. A549 cells were 100% authenticated with STR profiling. NMuMG and E14Tg2a.IV cells were validated with RNA-seq analysis for signature gene expression profile.

### Mycoplasma contamination

All cell lines tested negative for mycoplasma contamination.

### Commonly misidentified lines (See [ICLAC](#) register)

None of the cell lines used are listed as commonly misidentified lines in the ICLAC database.

## Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

### Laboratory animals

Pdx1-Cre; Cdkn2a  $-/-$ ;LSL-YFP (CIY) and Pdx1-Cre; KrasG12D $+/+$ ; Cdkn2a  $-/-$ ; Smad4  $-/-$  (KSIC) mice were provided by Nabeel Bardeesy. FVB/NJ (strain 001800) and B6129SF1/J (strain 101043) mouse strains were obtained from the Jackson Laboratory. Athymic nude mice (Hsd:Athymic Nude-Foxn1nu, 069) were obtained from Envigo. For all cancer cell injection studies, female mice were used between ages 5-7 weeks of age. All animal experiments were conducted in accordance with protocols approved by the MSKCC Institutional Animal Care and Use Committee and were in compliance with the relevant ethical regulations regarding animal research.

### Wild animals

The study did not involve wild animals.

### Field-collected samples

The study did not involve samples collected in the field.

## ChIP-seq

## Data deposition

- ☒ Confirm that both raw and final processed data have been deposited in a public database such as [GEO](#).
- ☒ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

## Data access links

*May remain private before publication.*

<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE118765>

The following secure token has been created to allow review of record GSE118765 while it remains in private status: ersrqismjhypxex.

## Files in database submission

## RAW FILES:

05\_806\_S4\_Input\_R1.combine.fastq.gz  
 06\_806\_S4\_SB\_SMAD23\_R1.combine.fastq.gz  
 07\_806\_S4\_TGFb\_S23\_R1.combine.fastq.gz  
 08\_806\_S4\_RREB1\_Input\_R1.combine.fastq.gz  
 09\_806\_S4\_RREB1\_SB\_HA\_R1.combine.fastq.gz  
 10\_806\_S4\_RREB1\_TGFb\_HA\_R1.combine.fastq.gz  
 11\_806\_S4\_RREB1\_WT\_Input\_R1\_001.fastq.gz  
 11\_806\_S4\_RREB1\_WT\_Input\_R2\_001.fastq.gz  
 12\_806\_S4\_RREB1\_WT\_SB\_SMAD23\_R1\_001.fastq.gz  
 12\_806\_S4\_RREB1\_WT\_SB\_SMAD23\_R2\_001.fastq.gz  
 13\_806\_S4\_RREB1\_WT\_TGFb\_SMAD23\_R1\_001.fastq.gz  
 13\_806\_S4\_RREB1\_WT\_TGFb\_SMAD23\_R2\_001.fastq.gz  
 14\_806\_S4\_RREB1\_KO\_Input\_R1\_001.fastq.gz  
 14\_806\_S4\_RREB1\_KO\_Input\_R2\_001.fastq.gz  
 15\_806\_S4\_RREB1\_KO\_SB\_SMAD23\_R1\_001.fastq.gz  
 15\_806\_S4\_RREB1\_KO\_SB\_SMAD23\_R2\_001.fastq.gz  
 16\_806\_S4\_RREB1\_KO\_TGFb\_SMAD23\_R1\_001.fastq.gz  
 16\_806\_S4\_RREB1\_KO\_TGFb\_SMAD23\_R2\_001.fastq.gz  
 33\_EB\_D3\_Input\_R1\_001.fastq.gz  
 33\_EB\_D3\_Input\_R2\_001.fastq.gz  
 34\_EB\_D3\_SMAD23\_R1\_001.fastq.gz  
 34\_EB\_D3\_SMAD23\_R2\_001.fastq.gz  
 35\_EB\_D3\_RREB1\_R1\_001.fastq.gz  
 35\_EB\_D3\_RREB1\_R2\_001.fastq.gz  
 36\_ATAC\_EB\_D3\_R1\_001.fastq.gz  
 36\_ATAC\_EB\_D3\_R2\_001.fastq.gz

## PROCESSED DATA FILES:

02\_806\_S4\_Input\_Tags.ucsc.bedGraph.gz  
 03\_806\_S4\_SB\_SMAD23\_Tags.ucsc.bedGraph.gz  
 04\_806\_S4\_TGFb\_SMAD23\_Tags.ucsc.bedGraph.gz  
 05\_806\_S4\_RREB1\_Input\_Tags.ucsc.bedGraph.gz  
 06\_806\_S4\_RREB1\_SB\_HA\_Tags.ucsc.bedGraph.gz  
 07\_806\_S4\_RREB1\_TGFb\_HA\_Tags.ucsc.bedGraph.gz  
 08\_806\_S4\_RREB1\_WT\_Input.sam\_Tags.ucsc.bedGraph.gz  
 09\_806\_S4\_RREB1\_WT\_SB\_SMAD23.sam\_Tags.ucsc.bedGraph.gz  
 10\_806\_S4\_RREB1\_WT\_TGFb\_SMAD23.sam\_Tags.ucsc.bedGraph.gz  
 11\_806\_S4\_RREB1\_KO\_Input.sam\_Tags.ucsc.bedGraph.gz  
 12\_806\_S4\_RREB1\_KO\_SB\_SMAD23.sam\_Tags.ucsc.bedGraph.gz  
 13\_806\_S4\_RREB1\_KO\_TGFb\_SMAD23.sam\_Tags.ucsc.bedGraph.gz  
 16\_EB\_D3\_Input\_Tags.ucsc.bedGraph.gz  
 17\_EB\_D3\_SMAD23\_Tags.ucsc.bedGraph.gz  
 18\_EB\_D3\_RREB1\_Tags.ucsc.bedGraph.gz  
 19\_ATAC\_EB\_D3\_Tags.ucsc.bedGraph.gz

## Genome browser session

(e.g. [UCSC](#))

[https://genome.ucsc.edu/cgi-bin/hgTracks?](https://genome.ucsc.edu/cgi-bin/hgTracks?hgS_doOtherUser=submit&hgS_otherUserName=suj&hgS_otherUserSessionName=Submission)

hgS\_doOtherUser=submit&hgS\_otherUserName=suj&hgS\_otherUserSessionName=Submission

## Methodology

## Replicates

All ChIP-seq experiments were performed 1 or 2 times, and confirmed with ChIP-PCR experiment.

## Sequencing depth

50 bp single-end or paired-end sequencing was performed to obtain 30 million read depth

## Antibodies

Rabbit monoclonal anti-SMAD2/3 Cell Signaling Technology Cat #8685; Rabbit monoclonal anti-HA-Tag Cell Signaling Technology Cat #3724.

## Peak calling parameters

Peak calling from ChIP-Seq data was performed with MACS 1.4.2 and verified by HOMER (v4.10). The parameters for peak calling included fold change >8, p value < 1e-8 to detect high confidence binding events. Input samples were used as reference controls for background correction.

Data quality

Details of data analysis and quality assurance are in the methods section

Software

For QC: FastQC v0.11.5  
For alignment: Bowtie2  
For ChIPseq sorting, normalization, and visualization: Sam Tools & HOMER



# An anti-CRISPR viral ring nuclease subverts type III CRISPR immunity

<https://doi.org/10.1038/s41586-019-1909-5>

Received: 24 July 2019

Accepted: 14 November 2019

Published online: 15 January 2020

Januka S. Athukoralage<sup>1</sup>, Stephen A. McMahon<sup>1</sup>, Changyi Zhang<sup>3,4</sup>, Sabine Grischow<sup>1</sup>, Shirley Graham<sup>1</sup>, Mart Krupovic<sup>2</sup>, Rachel J. Whitaker<sup>3,4</sup>, Tracey M. Gloster<sup>1\*</sup> & Malcolm F. White<sup>1\*</sup>

The CRISPR system in bacteria and archaea provides adaptive immunity against mobile genetic elements. Type III CRISPR systems detect viral RNA, resulting in the activation of two regions of the Cas10 protein: an HD nuclease domain (which degrades viral DNA)<sup>1,2</sup> and a cyclase domain (which synthesizes cyclic oligoadenylates from ATP)<sup>3–5</sup>. Cyclic oligoadenylates in turn activate defence enzymes with a CRISPR-associated Rossmann fold domain<sup>6</sup>, sculpting a powerful antiviral response<sup>7–10</sup> that can drive viruses to extinction<sup>7,8</sup>. Cyclic nucleotides are increasingly implicated in host–pathogen interactions<sup>11–13</sup>. Here we identify a new family of viral anti-CRISPR (Acr) enzymes that rapidly degrade cyclic tetra-adenylate (cA<sub>4</sub>). The viral ring nuclease AcrIII-1 is widely distributed in archaeal and bacterial viruses and in proviruses. The enzyme uses a previously unknown fold to bind cA<sub>4</sub> specifically, and a conserved active site to rapidly cleave this signalling molecule, allowing viruses to neutralize the type III CRISPR defence system. The AcrIII-1 family has a broad host range, as it targets cA<sub>4</sub> signalling molecules rather than specific CRISPR effector proteins. Our findings highlight the crucial role of cyclic nucleotide signalling in the conflict between viruses and their hosts.

Previously, we identified in the archaeon *Sulfolobus solfataricus* a family of cellular enzymes—referred to hereafter as the CRISPR-associated ring nuclease I (CrnI) family—that degrades cA<sub>4</sub> molecules and deactivates the cA<sub>4</sub>-dependent RNase Csx1<sup>14</sup>. This enzyme is thought to act by mopping up cA<sub>4</sub> molecules in the cell without compromising the immunity provided by the type III CRISPR system. In the absence of such a mechanism to remove cyclic oligoadenylates (cOAs) following the clearance of viral infections, cells could be pushed towards dormancy or cell death under inappropriate circumstances<sup>7,14</sup>. Unsurprisingly, viruses have responded to the threat of the CRISPR system by evolving a range of anti-CRISPR (Acr) proteins, which are used to inhibit and overcome the cell's CRISPR defences using a variety of mechanisms (reviewed in ref. <sup>15</sup>). Acrs have been identified for many of the CRISPR effector subtypes, and number more than 40 families<sup>16</sup>.

Here we investigate the DUF1874 protein family, which is conserved and widespread in a variety of archaeal viruses and plasmids, bacteriophages and prophages (Extended Data Fig. 1), for an Acr function. Structures are available for several members of the DUF1874 family, including gp29 from *Sulfolobus islandicus* rod-shaped virus 1 (SIRV1)<sup>17</sup> and B116 from *Sulfolobus* turreted icosahedral virus (STIV)<sup>18</sup>. The structures reveal an intriguing dimeric structure, with a large central pocket flanked by conserved residues. B116 is also known to be important for normal virus replication kinetics, as deletion of the gene results in a marked 'small plaque' phenotype<sup>19</sup>, consistent with an Acr function.

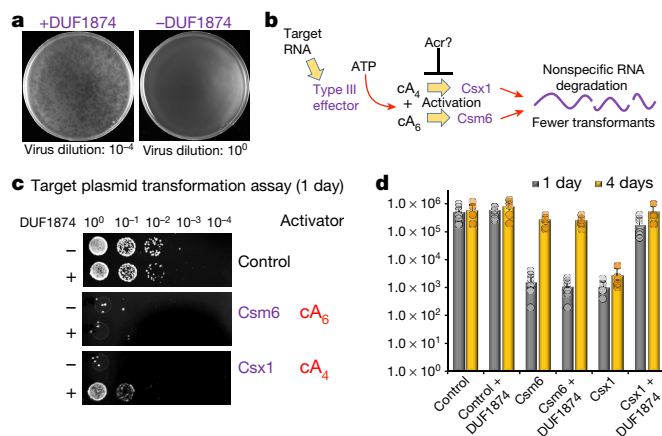
## DUF1874 is a type III anti-CRISPR, AcrIII-1

To investigate a possible Acr function of DUF1874, we deleted the genes for the type I-A CRISPR system in *Sulfolobus islandicus* M.16.4, so that it had only a type III-B system for defence<sup>20</sup> (Extended Data Fig. 2). We challenged this strain with the archaeal virus SSeV (Fig. 1a), a lytic virus isolated from Kamchatka, Russia, that has an exact CRISPR-spacer match of 100% in M.16.4, as well as several other potentially active CRISPR spacers. SSeV lacks a *duf1874* gene and failed to form plaques on a lawn of *S. islandicus* M.16.4 with type III-B CRISPR defence unless the effector gene *csx1* was deleted (Fig. 1a and Extended Data Fig. 2). However, the same cells expressing the SIRV1 gp29 gene from a plasmid were readily infected, giving rise to plaque formation. These data are consistent with the hypothesis that SIRV1 gp29 functions as an Acr specific for the type III CRISPR defence.

To explore this possibility further, we used a recently developed recombinant type III CRISPR system from *Mycobacterium tuberculosis*; this system allows the effector protein downstream of cOAs to be swapped in order to provide effective immunity based on either cA<sub>6</sub> or cA<sub>4</sub> signalling<sup>21</sup> (Fig. 1). We then transformed strains capable of cA<sub>4</sub>- or cA<sub>6</sub>-based immunity with a plasmid that was targeted for interference owing to a match in its tetracycline-resistance gene to a spacer in the CRISPR array. We observed efficient interference (lack of plasmid transformation) after one day for either strain in the absence of the *duf1874* gene from bacteriophage THSA-485A (Fig. 1c,d). However, the presence of the *duf1874* gene on the

<sup>1</sup>Biomedical Sciences Research Complex, School of Biology, University of St Andrews, St Andrews, Fife, UK. <sup>2</sup>Department of Microbiology, Institut Pasteur, Paris, France. <sup>3</sup>Department of Microbiology, University of Illinois at Urbana-Champaign, Urbana, IL, USA. <sup>4</sup>Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL, USA.

\*e-mail: tmg@st-andrews.ac.uk; mfw2@st-andrews.ac.uk

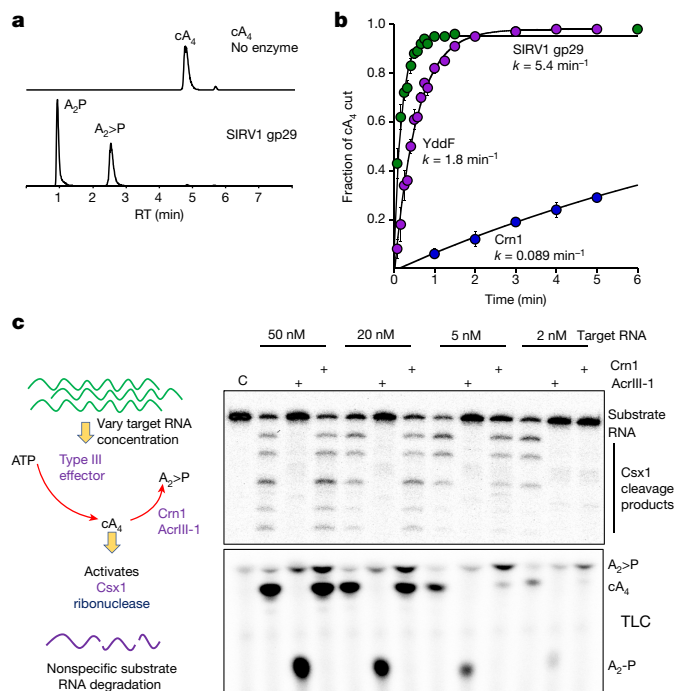


**Fig. 1 | DUF1874 is an anti-CRISPR protein specific for  $cA_4$  signalling.** **a**, SSeV infection assay, showing that *gp29* (a *duf1874* gene from SIRV1) can neutralize the type III-B CRISPR system in *S. islandicus*. We challenged *S. islandicus* RJW007Δtype I-A or RJW007Δtype I-AΔ*csx1* mutant strains with SSeV, in the presence or absence of *duf1874* (SIRV1 *gp29*) expressed on a replicative plasmid. Plaques were observed when *csx1* was deleted, or when the resistant strain expressed *duf1874* ( $n = 3$  biological replicates) (Extended Data Fig. 2d). **b**, Diagram showing the reconstituted *M. tuberculosis* type III-A CRISPR interference system established in *E. coli*. By swapping the native ancillary nuclease Csm6 for Csx1, the system can be converted from  $cA_6$ - to  $cA_4$ -mediated antiviral immunity. **c**, Plasmid transformation assay (after one day's growth), using a plasmid with a match to a spacer in the CRISPR array. If the plasmid is successfully targeted by the CRISPR system, fewer transformants are expected. Plasmids with or without the *duf1874* gene were targeted successfully when  $cA_6$  (Csm6)-mediated antiviral signalling was active. By contrast, cells using a  $cA_4$  (Csx1)-based system reduced transformation only when *duf1874* was not present, suggesting that DUF1874 was effective in neutralizing  $cA_4$ -based CRISPR interference. The control strain lacked *cOA*-dependent ribonucleases. These results are representative of two biological replicates, with four technical replicates each ( $n = 8$ ). **d**, Colony counts for transformants visible after one and four days' growth in the presence or absence of DUF1874 and the indicated effector proteins. DUF1874 antagonizes Csx1- but not Csm6-mediated immunity. Data are mean and s.d. from two biological replicates with four technical replicates each ( $n = 8$ ).

plasmid reduced immunity for  $cA_4$ -mediated, but not  $cA_6$ -mediated, CRISPR defence. This observation supports the hypothesis that DUF1874 acts as an Acr against  $cA_4$ -mediated type III CRISPR defence. We therefore propose the collective name AcrIII-1 for this family. The '-' in place of the subtype reflects the fact that AcrIII-1 will inhibit any type III CRISPR subtype that utilizes  $cA_4$  molecules for defence<sup>22</sup>. We also found that, after four days of growth, Csm6-mediated immunity was lost, regardless of the presence of DUF1874. This could indicate that alternative mechanisms exist to remove  $cA_6$  (Fig. 1d and Extended Data Fig. 3).

### AcrIII-1 degrades $cA_4$ rapidly

To explore the mechanism of action of the AcrIII-1 family, we cloned and expressed two family members in *Escherichia coli*: the SIRV1 *gp29* protein and the YddF protein, encoded by an integrative and conjugative element (ICE), *Bsl*, from *Bacillus subtilis*<sup>23</sup> (Extended Data Fig. 1b). We found that both proteins possess potent ring nuclease activity, rapidly degrading  $cA_4$  to generate linear di-adenylate (ApA>P) with a cyclic 2',3'-phosphate (Fig. 2 and Extended Data Fig. 4). With a catalytic rate exceeding  $5 \text{ min}^{-1}$ , the Acr enzyme is at least 60-fold more active than the cellular ring nuclease Crn1 from *S. solfataricus*. Both SIRV1 *gp29* and YddF show a strong preference for  $cA_4$  over  $cA_6$ , with the latter being degraded very slowly by comparison (Extended Data Fig. 4). We showed previously that the type III-D CRISPR effector of *S. solfataricus* generates  $cA_4$  in proportion to the amount of cognate target RNA

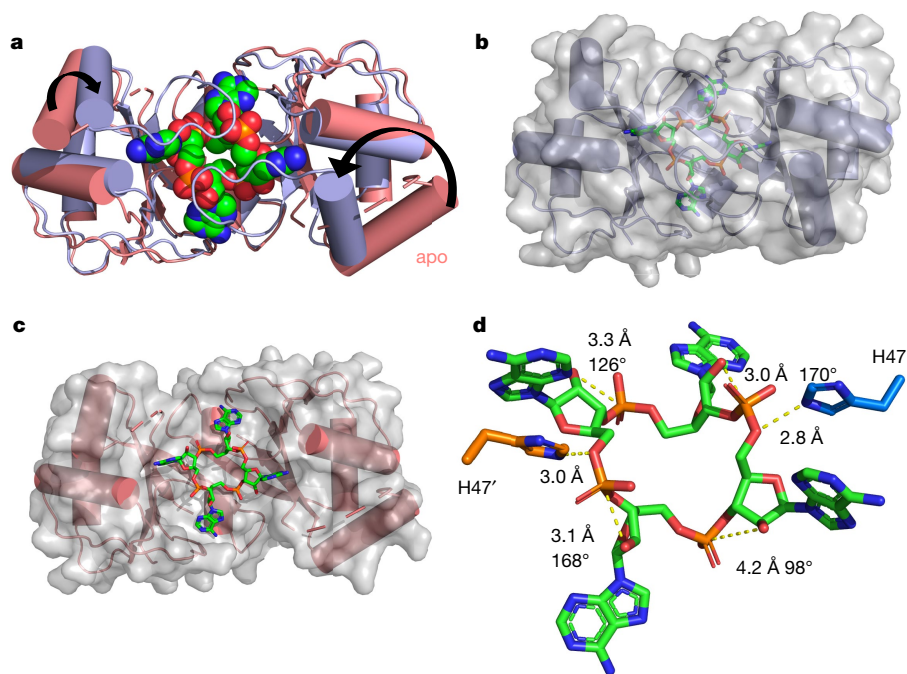


**Fig. 2 | AcrIII-1 rapidly degrades  $cA_4$  to linear products.** **a**, Liquid chromatography/high-resolution mass spectrometry analysis confirms that AcrIII-1 SIRV1 *gp29* converts  $cA_4$  to  $A_2>P$  and  $A_2-P$ . The experiment was repeated twice with similar results. RT, retention time. **b**, Kinetic comparison of  $cA_4$  degradation by the AcrIII-1 enzymes SIRV1 *gp29* and YddF and the cellular ring nuclease Crn1. Values and error bars reflect means  $\pm$  standard deviation ( $n = 3$  technical replicates). **c**, The left panel shows the experimental protocol. On the right, the top panel shows activation of Csx1 in a coupled assay containing type III Csm complex activated with the indicated amounts of (unlabelled) target RNA to initiate  $cA_4$  synthesis. The control (C) reaction comprises Csx1 and substrate RNA alone. Each set of three lanes thereafter is first in the absence and then in the presence of a Crn1 protein (Sso2081) or an AcrIII-1 protein (SIRV1 *gp29*). Whereas AcrIII-1 degraded all  $cA_4$  molecules generated using up to 50 nM of the target RNA, the Crn1 enzyme deactivated Csx1 only when less than 5 nM RNA was used. The bottom panel shows thin layer chromatography (TLC) of the same reactions to visualize  $cA_4$  production and degradation. Csx1 deactivation correlated with complete  $cA_4$  degradation ( $n = 3$  technical replicates). For gel source data, see Supplementary Fig. 1.

present<sup>14</sup>. By varying the target RNA input and following  $cA_4$  levels and Csx1 activity, we compared the abilities of Crn1 and AcrIII-1 to destroy the signalling molecule and deactivate the ancillary defence nuclease Csx1. In keeping with its low turnover number, Crn1 was effective at degrading  $cA_4$  and thus deactivating Csx1 only at the lowest levels of target RNA (Fig. 2c). By contrast, AcrIII-1 degraded  $cA_4$  completely at the highest target RNA concentration examined, preventing Csx1 activation. We investigated the ability of each enzyme to prevent Csx1 activation over a range of  $cA_4$  concentrations spanning four orders of magnitude (Extended Data Fig. 4e). Crn1 (2  $\mu\text{M}$ ) provided protection only up to 5  $\mu\text{M}$   $cA_4$ , but 2  $\mu\text{M}$  of AcrIII-1 provided complete protection at the highest level of  $cA_4$  tested (500  $\mu\text{M}$ ). Thus, AcrIII-1 has the potential to destroy large concentrations of the second messenger  $cA_4$  rapidly, preventing activation of Csx1.

### Structure and mechanism of AcrIII-1

The structure of AcrIII-1 is unrelated to that of proteins with the CRISPR-associated Rossmann fold (CARF) domain—the only protein family known thus far to bind *cOA*<sup>6</sup>. To elucidate the mechanism of  $cA_4$  binding and cleavage by AcrIII-1, we co-crystallized an inactive variant (H47A) of SIRV1 *gp29* with  $cA_4$ , and solved the structure to 1.55 Å resolution



**Fig. 3 | Structure of AcrIII-1 bound to cA<sub>4</sub>.** **a**, Superimposition of the apo SIRV1 gp29 structure (salmon) and the same protein in complex with cA<sub>4</sub> (purple), highlighting the movement of the loop and  $\alpha$ -helix upon cA<sub>4</sub> binding. cA<sub>4</sub> is shown coloured by element. **b**, Surface representation of the structure of SIRV1 gp29 (purple) in complex with cA<sub>4</sub>, emphasizing the complete burial of the ligand. **c**, Surface representation of the apo structure of SIRV1 gp29 (salmon) with cA<sub>4</sub> in the position observed in the structure of the complex, indicating that the binding site is preformed. **d**, Structure of cA<sub>4</sub> bound to SIRV1 gp29. The

two active-site histidine residues (H47A and H47A', from each monomer of the dimer; modelled on the basis of the position of the alanine side chain in the H47A variant crystallized with cA<sub>4</sub>, and coloured to represent residues from different monomers) are in suitable positions to act as the general acid, protonating the oxyanion leaving group. The corresponding ribose sugars have 2'-hydroxyl groups suitably positioned for in-line nucleophilic attack on the phosphodiester bond. In the cA<sub>4</sub> ligand, carbon atoms are shown in green, phosphates in orange, oxygens in red and nitrogens in blue.

(Fig. 3 and Extended Data Table 1). The complex reveals a molecule of cA<sub>4</sub> bound at the dimer interface. Comparison of the cA<sub>4</sub>-bound and apo structures reveals a substantial movement of a loop (comprising residues 82–94) and subsequent  $\alpha$ -helix to bury cA<sub>4</sub> within the dimer. These loops adopt variable or unstructured conformations in the various apo protein structures. Once bound, the ligand is completely enclosed by the protein—a considerable accomplishment when one considers the relative sizes of protein and ligand (Fig. 3b). Superimposition of the cA<sub>4</sub> ligand on the apo-protein structure reveals that the binding site is largely preformed, with the exception of the mobile loops that form the lid (Fig. 3c). The overall change is like two cupped hands catching a ball, with the loops (fingers) subsequently closing around it.

The cA<sub>4</sub> molecule makes symmetrical interactions with each monomer of AcrIII-1 (Extended Data Fig. 5). Arginine R85 on the loop from one monomer interacts with the distant half of the cA<sub>4</sub> molecule and appears to 'lock' the closed dimer. Other important interactions are made with main-chain L92, I69 and N8, and side-chain R66, N8, Q81, S11, T50, S49 and N13, most of which are semi or fully conserved (Extended Data Figs. 1, 5), suggesting that they have important roles in cA<sub>4</sub> binding and/or catalysis in this whole family of enzymes. At two positions, on opposite sides of the ring, the 2'-hydroxyl of the ribose is positioned correctly for in-line attack on the phosphodiester bond, consistent with the observed bilateral cleavage (Fig. 3d). The catalytic power of the AcrIII-1 family probably derives from active-site residues that position the 2'-hydroxyl group for in-line nucleophilic attack, stabilize the transition state and protonate the oxyanion leaving group<sup>24</sup>. For the AcrIII-1 family, the absolutely conserved residue H47 is suitably positioned to act as a general acid and fulfil the latter role (Fig. 3d). To test this hypothesis, we assayed variant H47A of AcrIII-1. The variant enzyme suffered a more than 2,500-fold decrease in catalytic power, which could be partially reversed by chemical rescue with 500 mM imidazole in the reaction buffer (Extended Data Fig. 6). We also noted that

the conserved residue E88, situated on the tip of the loop that covers the binding site, is positioned close to the H47 residue of the opposite subunit. When mutated to alanine, the catalytic rate was reduced by 84-fold to 0.064 min<sup>-1</sup> (Extended Data Fig. 6b), consistent with a role for E88 in positioning H47 and/or increasing the pK<sub>a</sub> of the catalytic histidine residue to enhance catalysis<sup>25</sup>.

By targeting a key signalling molecule, a single AcrIII-1 enzyme should have broad utility in the inhibition of endogenous cA<sub>4</sub>-specific type III CRISPR systems in any species. Of the CRISPR ancillary nucleases studied to date, most are activated by cA<sub>4</sub>; activation by cA<sub>6</sub> appears to be limited to certain bacterial phyla, including the Firmicutes and Actinobacteria<sup>21</sup>. Recently, a type III Acr (AcrIIIB1) has been reported that appears to function by binding and inhibiting the type III-B effector complex<sup>26</sup>. Two other Acr proteins with enzymatic functions have been described: AcrVA1, which catalyses CRISPR RNA (crRNA)-mediated cleavage of Cas12a<sup>27</sup>, and AcrVA5, which acetylates the site in Cas12a that senses the protospacer-adjacent motif (PAM) of target DNA<sup>28</sup>. These and other Acrs target a protein (or protein/nucleic acid complex), implying a requirement for specific interactions that could be evaded by sequence variation. This is not a limitation of AcrIII-1.

### Phylogenetic analysis of AcrIII-1

The gene encoding AcrIII-1 is found in representatives of at least five distinct viral families, making it one of the most widely conserved of all archaeal virus proteins<sup>29</sup> (Extended Data Fig. 1 and Supplementary Data 1). The distribution of AcrIII-1 in archaea is sporadic but covers most of the main lineages (Supplementary Data 1), and is typically adjacent to open reading frames (ORFs) from mobile genetic elements rather than CRISPR loci. A good example is the STIV integrated into *S. acidocaldarius* genomes<sup>30</sup>. AcrIII-1 is also present in several bacteriophages of the order *Caudovirales*, and there are many instances of *acrIII-1* genes

in sequenced bacterial genomes, with homologues found in the Firmicutes, Cyanobacteria, Proteobacteria, Actinobacteria and other phyla (Supplementary Data 1). Maximum likelihood phylogenetic analysis of the *AcrIII-1* proteins suggests multiple horizontal gene transfers between unrelated viruses, as well as between bacteria and archaea (Extended Data Figs. 7–9). Sometimes the *acrIII-1* gene is clearly part of an integrated mobile genetic element, such as the *yddF* gene in *B. subtilis*<sup>29</sup>. However, in other species ( $n = 49$ ) the gene is associated with cellular type III CRISPR systems. In *Marinitoga piezophila*, *AcrIII-1* is fused to a cOA-activated HEPN RNase of the Csx1 family. Given that both active sites are conserved, this fusion protein may have cA<sub>4</sub>-activated RNase activity coupled with a cA<sub>4</sub>-degradative ring nuclease, thus providing an explicit linkage between the *AcrIII-1* family and the type III CRISPR system. In this context the enzyme is likely to be acting as a host-encoded ring nuclease, like *Crn1*, rather than an *Acr*. We therefore propose the family name of *Crn2* (CRISPR-associated ring nuclease 2) to cover DUF1874-family members that are associated with type III CRISPR systems (Extended Data Fig. 8).

## Cyclic nucleotides in defence systems

*AcrIII-1* is, to our knowledge, the first *Acr* to be predicted to have functional roles in both ‘offence and defence’. It remains to be determined whether the *acrIII-1* gene arose in viruses and was appropriated by cellular type III systems, or vice versa. However, the extremely broad distribution of *acrIII-1* and limited distribution of *crn2* suggests the former. Adoption of an anti-CRISPR protein as a component of a cellular CRISPR defence system seems counterintuitive. However, the enzyme could have been harnessed for a role in defence by putting it under tight transcriptional control so that it is expressed at appropriate times or levels. The unprecedentedly wide occurrence of this *Acr* across many archaeal and bacterial virus families reflects the fact that this enzyme degrades a key signalling molecule to subvert cellular immunity. This makes it very hard for cells to evolve counter-resistance, other than by switching to a different signalling molecule. Recent discoveries have highlighted the existence of diverse cellular defence systems involving cyclic nucleotide signalling in bacteria<sup>11–13</sup>. It is possible that cOAs, and the enzymes that metabolize them, have functions that extend beyond type III CRISPR systems. The identification here of a new class of cA<sub>4</sub>-binding proteins highlights the potential for further discoveries in this area.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-019-1909-5>.

1. Samai, P. et al. Co-transcriptional DNA and RNA cleavage during type III CRISPR-Cas immunity. *Cell* **161**, 1164–1174 (2015).
2. Tamulaitis, G. et al. Programmable RNA shredding by the type III-A CRISPR-Cas system of *Streptococcus thermophilus*. *Mol. Cell* **56**, 506–517 (2014).

3. Kazlauskienė, M., Kostiuk, G., Venclovas, Č., Tamulaitis, G. & Siksnys, V. A cyclic oligonucleotide signaling pathway in type III CRISPR-Cas systems. *Science* **357**, 605–609 (2017).
4. Niewoehner, O. et al. Type III CRISPR-Cas systems produce cyclic oligoadenylate second messengers. *Nature* **548**, 543–548 (2017).
5. Rouillon, C., Athukoralage, J. S., Graham, S., Grischow, S. & White, M. F. Control of cyclic oligoadenylate synthesis in a type III CRISPR system. *eLife* **7**, e36734 (2018).
6. Makarova, K. S., Anantharaman, V., Grishin, N. V., Koonin, E. V. & Aravind, L. CARF and WYL domains: ligand-binding regulators of prokaryotic defense systems. *Front. Genet.* **5**, 102 (2014).
7. Rostøl, J. T. & Marraffini, L. A. Non-specific degradation of transcripts promotes plasmid clearance during type III-A CRISPR-Cas immunity. *Nat. Microbiol.* **4**, 656–662 (2019).
8. Pyenson, N. C., Gayvert, K., Varble, A., Elemento, O. & Marraffini, L. A. Broad targeting specificity during bacterial type III CRISPR-Cas immunity constrains viral escape. *Cell Host Microbe* **22**, 343–353 (2017).
9. Deng, L., Garrett, R. A., Shah, S. A., Peng, X. & She, Q. A novel interference mechanism by a type IIIB CRISPR-Cmr module in *Sulfolobus*. *Mol. Microbiol.* **87**, 1088–1099 (2013).
10. Jiang, W., Samai, P. & Marraffini, L. A. Degradation of phage transcripts by CRISPR-associated RNases enables type III CRISPR-Cas immunity. *Cell* **164**, 710–721 (2016).
11. Whiteley, A. T. et al. Bacterial cGAS-like enzymes synthesize diverse nucleotide signals. *Nature* **567**, 194–199 (2019).
12. Maelfait, J. & Rehwinkel, J. RECONsidering sensing of cyclic dinucleotides. *Immunity* **46**, 337–339 (2017).
13. Cohen, D. et al. Cyclic GMP-AMP signalling protects bacteria against viral infection. *Nature* **574**, 691–695 (2019).
14. Athukoralage, J. S., Rouillon, C., Graham, S., Grischow, S. & White, M. F. Ring nucleases deactivate type III CRISPR ribonucleases by degrading cyclic oligoadenylate. *Nature* **562**, 277–280 (2018).
15. Borges, A. L., Davidson, A. R. & Bondy-Denomy, J. The discovery, mechanisms, and evolutionary impact of anti-CRISPRs. *Annu. Rev. Virol.* **4**, 37–59 (2017).
16. Hwang, S. & Maxwell, K. L. Meet the anti-CRISPRs: widespread protein inhibitors of CRISPR-Cas systems. *CRISPR J.* **2**, 23–30 (2019).
17. Oke, M. et al. The Scottish structural proteomics facility: targets, methods and outputs. *J. Struct. Funct. Genomics* **11**, 167–180 (2010).
18. Larson, E. T. et al. A new DNA binding protein highly conserved in diverse crenarchaeal viruses. *Virology* **363**, 387–396 (2007).
19. Wirth, J. F. et al. Development of a genetic system for the archaeal virus *Sulfolobus turreted icosahedral virus* (STIV). *Virology* **415**, 6–11 (2011).
20. Bautista, M. A., Zhang, C. & Whitaker, R. J. Virus-induced dormancy in the archaeon *Sulfolobus islandicus*. *MBio* **6**, e02565-14 (2015).
21. Grischow, S., Athukoralage, J. S., Graham, S., Hoogeboom, T. & White, M. F. Cyclic oligoadenylate signalling mediates *Mycobacterium tuberculosis* CRISPR defence. *Nucleic Acids Res.* **47**, 9259–9270 (2019).
22. Bondy-Denomy, J. et al. A unified resource for tracking anti-CRISPR names. *CRISPR J.* **1**, 304–305 (2018).
23. Auchtung, J. M., Aleksanyan, N., Bulku, A. & Berkmen, M. B. Biology of ICEBs1, an integrative and conjugative element in *Bacillus subtilis*. *Plasmid* **86**, 14–25 (2016).
24. Yang, W. Nucleases: diversity of structure, function and mechanism. *Q. Rev. Biophys.* **44**, 1–93 (2011).
25. Broo, K. S., Brive, L., Sott, R. S. & Baltzer, L. Site-selective control of the reactivity of surface-exposed histidine residues in designed four-helix-bundle catalysts. *Fold. Des.* **3**, 303–312 (1998).
26. Bhoobalan-Chitty, Y., Johansen, T. B., Di Cianni, N. & Peng, X. Inhibition of type III CRISPR-Cas immunity by an archaeal virus-encoded anti-CRISPR protein. *Cell* **179**, 448–458 (2019).
27. Knott, G. J. et al. Broad-spectrum enzymatic inhibition of CRISPR-Cas12a. *Nat. Struct. Mol. Biol.* **26**, 315–321 (2019).
28. Dong, L. et al. An anti-CRISPR protein disables type V Cas12a by acetylation. *Nat. Struct. Mol. Biol.* **26**, 308–314 (2019).
29. Keller, J. et al. Crystal structure of AFV3-109, a highly conserved protein from crenarchaeal viruses. *Virol. J.* **4**, 12 (2007).
30. Anderson, R. E., Kouris, A., Seward, C. H., Campbell, K. M. & Whitaker, R. J. Structured populations of *Sulfolobus acidocaldarius* with susceptibility to mobile genetic elements. *Genome Biol. Evol.* **9**, 1699–1710 (2017).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2020



## Methods

### Construction of *S. islandicus* strains

The type I-A CRISPR defence module, which includes seven genes—*cas3b*, *cas5*, *cas7*, *cas5*, *cas3'*, *cas3''* and *casX*<sup>31</sup>—was in-frame deleted from the genetic host *S. islandicus* RJW007, derived from wild-type strain *S. islandicus* M.16.4 carrying a double *pyrEF* and *argD* deletion<sup>32</sup>, by using a modified plasmid integration and segregation knockout strategy<sup>33</sup>, in line with the methodology developed in ref.<sup>34</sup>. The resultant type I-A deletion mutant (RJW007Δtype I-A) was then used as a parental strain to further delete the *csxI* gene, generating the mutant strain RJW007Δtype I-AΔ*csxI*. Mutant strains were confirmed by polymerase chain reaction (PCR) analysis using primers that bind outside of the homologous flanking arms of genes to be deleted.

Synthesized SIRV1 *gp29* gene was purchased from Integrated DNA Technologies (IDT), Coralville, USA as a g-block, and was cloned into a *Sulfolobus*-*E. coli* shuttle vector, pSeSd-SsoargD<sup>32</sup> (referred to as pOE hereafter), at the NdeI and NotI sites, generating the *gp29* expression plasmid pOE-*gp29* in which the *gp29* gene was placed under the control of the arabinose promoter. The pOE-*gp29* and pOE plasmids were then transformed into competent cells of the Δtype I-A mutant and Δtype I-AΔ*csxI* mutant via electroporation as described<sup>32</sup>, generating strains expressing and not expressing SIRV1 *gp29*, respectively.

### Viral quantification

The genome sequence of SSeV is available in GenBank (accession code MN53972). To calculate the titre of SSeV, we co-incubated 100 µl diluted virus ( $10^{-5}$ ,  $10^{-6}$  and  $10^{-7}$ ) with 500 µl *S. islandicus* Y08.82.36 host<sup>20</sup> (ten-fold concentrated) without shaking at 76–78 °C for 30 min. Afterwards, we transferred the virus-infected cells into a glass test tube containing 5 ml of prewarmed sucrose-yeast extract (SY) and 0.8% gelrite mixture, and plated onto SY plates. The plates were put into a plastic bag, and incubated for two days at 76–78 °C. We counted plaques in plates with proper virus dilutions, and determined the titre of SSeV to be  $4.96 \times 10^8$  plaque-forming units (PFUs) per millilitre.

### SSeV infection of *S. islandicus* with or without type III CRISPR

We carried out the SSeV infection assay as described<sup>20</sup>, with minor modifications. In brief, approximately  $6 \times 10^8$  cells of *S. islandicus* M.16.4 cells taken from the exponential stage were spun down at 4,000 r.p.m. for 12 min, and resuspended in 1 ml of arabinose-tryptone (AT) medium. The resuspensions were then co-incubated with 20 ml of fresh AT medium or SSeV supernatant at different dilutions ( $10^0$ ,  $10^{-1}$ ,  $10^{-2}$ ,  $10^{-3}$ ,  $10^{-4}$ ,  $10^{-5}$  and  $10^{-6}$ ) in a Falcon tube at 76–78 °C for 1 h without shaking. The SSeV-infected cells were washed twice with 10 ml of AT medium and resuspended into 500 µl of AT medium. Afterwards, the concentrated SSeV-infected cells were mixed with 5 ml of top layer (2.5 ml of 2 × arabinose yeast extract (AY) medium plus 2.5 ml of 0.8% gelrite), and then plated onto AY plates. PFUs were counted after four days of incubation at 76–78 °C. Three independent experiments were performed.

### Cloning and purification

For cloning, we purchased synthetic genes (g-blocks) from IDT, and cloned them into the vector pEhisV5spacerTev between the NcoI and BamHI sites<sup>35</sup>. Competent DH5α (*E. coli*) cells were transformed with the construct, and sequence integrity confirmed by sequencing (Eurofins Genomics). The plasmid was transformed into *E. coli* C43 (DE3) cells for protein expression. Cloning of AcrIII-1 SIRV1 *gp29*, Crn1 Sso2081 and SsoC*csxI* has been described previously<sup>14,17</sup>. For expression of SIRV1 *gp29* and *Bacillus subtilis* YddF, we grew 2 l of Luria-Broth (LB) culture at 37 °C to an OD<sub>600</sub> of 0.8 with shaking at 180 r.p.m. Protein expression was induced with 0.4 mM isopropyl β-D-1-thiogalactopyranoside, and cells were grown at 25 °C overnight before harvesting by centrifugation at 4,000 r.p.m. (Beckman Coulter Avanti JXN-26; JLA8.1 rotor) at 4 °C for 15 min.

For protein purification, we resuspended the cell pellet in four volumes equivalent of lysis buffer containing 50 mM Tris-HCl 7.5, 0.5 M NaCl, 10 mM imidazole and 10% glycerol supplemented with EDTA-free protease-inhibitor tablets (Roche; one tablet per 100 ml buffer) and lysozyme (1 mg ml<sup>-1</sup>). Cells were lysed by sonicating six times for one minute with one-minute rest intervals on ice at 4 °C, and the lysate was ultracentrifuged at 40,000 r.p.m. (70 Ti rotor) at 4 °C for 35 min. The lysate was then loaded onto a 5 ml HisTrap FF Crude column (GE Healthcare) equilibrated with wash buffer containing 50 mM Tris-HCl pH 7.5, 0.5 M NaCl, 30 mM imidazole and 10% glycerol. Unbound protein was washed away with 20 column volumes of wash buffer, before elution of histidine-tagged protein using a linear gradient (holding at 10% for three column volumes, and 50% for three column volumes) of elution buffer containing 50 mM Tris-HCl pH 7.5, 0.5 M NaCl, 0.5 M imidazole and 10% glycerol. We carried out SDS-polyacrylamide gel electrophoresis (PAGE) to identify fractions containing the protein of interest, and pooled and concentrated relevant fractions using a 10 kDa molecular mass cut-off centrifugal concentrator (Merck). The histidine tag was removed by incubating concentrated protein overnight with tobacco etch virus (TEV) protease (1 mg per 10 mg protein) while dialysing in buffer containing 50 mM Tris-HCl pH 7.5, 0.5 M NaCl, 30 mM imidazole and 10% glycerol at room temperature. The protein with histidine tag removed was isolated using a 5 ml HisTrapFF column, eluting the protein using four column volumes of wash buffer. Histidine-tag-removed protein was further purified by size-exclusion chromatography (S200 16/60; GE Healthcare) in buffer containing 20 mM Tris-HCl pH 7.5, 0.125 M NaCl using an isocratic gradient. After SDS-PAGE, fractions containing protein of interest were concentrated and protein was aliquoted and stored at -80 °C. We generated variant enzymes using the QuickChange site-directed mutagenesis kit as per the manufacturer's instructions (Agilent Technologies), and purified them as for the wild-type proteins.

### Radiolabelled cA<sub>4</sub>-cleavage assays

We generated cOA by incubating 120 µg *Sulfolobus solfataricus* (Sso) type III-D (Csm) complex with 5 nM α-<sup>32</sup>P-ATP, 1 mM ATP, 120 nM A26 RNA target and 2 mM MgCl<sub>2</sub> in Csx1 buffer containing 20 mM 2-(*N*-morpholino)ethanesulfonic acid (MES) pH 5.5, 100 mM K-glutamate, 1 mM dithiothreitol (DTT) and three units SUPERase•In Inhibitor for 2 h at 70 °C in a 100 µl reaction volume. We extracted cOA through phenol-chloroform (Ambion) extraction followed by chloroform extraction (Sigma-Aldrich), with storage at -20 °C.

For single-turnover kinetics experiments, we assayed AcrIII-1 SIRV1 *gp29* and variants (4 µM protein dimer) for radiolabelled cA<sub>4</sub> degradation by incubating with 1/400 diluted <sup>32</sup>P-labelled SsoCsm cOA (roughly 200 nM cA<sub>4</sub>, generated in a 100 µl cOA-synthesis reaction as above) in Csx1 buffer supplemented with 1 mM EDTA at 50 °C. We incubated AcrIII-1 YddF (8 µM dimer) with cOA in buffer containing 20 mM MES pH 6.0, 100 mM NaCl, 1 mM DTT, 1 mM EDTA and three units SUPERase•In Inhibitor at 37 °C. We incubated Crn1 Sso2081 (4 µM dimer) with cOA in buffer containing 20 mM Tris-HCl pH 8.0, 100 mM NaCl, 1 mM EDTA, 1 mM DTT and three units SUPERase•In Inhibitor at 50 °C. For SIRV1 *gp29* H47A chemical rescue, reactions were supplemented with 0.5 M imidazole. Two experimenters were involved in kinetic experiments involving five-second time points. At desired time points, a 10 µl aliquot of the reaction was removed and quenched by adding to phenol chloroform and vortexing. Subsequently, 5 µl of deproteinized reaction product was extracted into 5 µl 100% formamide xylene-cyanol loading dye if intended for denaturing PAGE, or products were further isolated by chloroform extraction if intended for thin-layer chromatography (TLC). A reaction incubating cOA in buffer without protein to the endpoint of each experiment was included as a negative control. All experiments were carried out in triplicate. For SIRV1 *gp29*, two biological samples were assayed in triplicate. We visualized cA<sub>4</sub> degradation by phosphor imaging following denaturing PAGE (7 M urea, 20% acrylamide, 1× Tris/borate/EDTA (TBE)) or TLC.



For TLC, we spotted 1 µl of radiolabelled product 1 cm from the bottom of a 20 × 20 cm silica gel TLC plate with fluorescence indicator 254 nm (Supelco Sigma-Aldrich). We placed the TLC plate in a sealed glass chamber prewarmed and humidified at 37 °C and containing 0.5 cm of a running buffer composed of 30% water, 70% ethanol and 0.2 M ammonium bicarbonate, pH 9.2. The temperature was lowered to 35 °C and the buffer was allowed to rise along the plate through capillary action until the migration front reached 17 cm. The plate was air dried and sample migration was visualized by phosphor imaging.

To examine degradation of cA<sub>4</sub> and cA<sub>6</sub> by AcrIII-1 proteins, we incubated unlabelled cA<sub>4</sub> or cA<sub>6</sub> (450 µM, BIOLOG Life Science Institute, Bremen, Germany) with SIRV1 gp29 or YddF (40 µM dimer), in reaction buffers described above, at 70 °C and 37 °C, respectively. Reactions were quenched at the indicated time points and prepared for TLC as above. We visualized reaction substrate and products, which block fluorescence of the indicator on the plate, under shortwave UV light (254 nm) and photographed the plates using a 12-megapixel/1.8-aperture camera.

For kinetic analysis, we quantified cA<sub>4</sub> cleavage using the Bio-Formats plugin<sup>36</sup> of ImageJ as distributed in the Fiji package<sup>37</sup> and fitted the data to a single exponential curve ( $y = m1 + m2 \cdot (1 - \exp(-m3 \cdot x))$ ;  $m1 = 0.1$ ,  $m2 = 1$  and  $m3 = 1$ ) using Kaleidagraph (Synergy Software), as before<sup>38</sup>. We obtained the cA<sub>4</sub>-cleavage rate by the H47A variant in the absence of imidazole by linear fit. Raw data for kinetic analyses are available in Supplementary Data 2.

### Deactivation of HEPN nucleases by ring nucleases

In the absence or presence of Crn1 Sso2081 (2 µM dimer) or AcrIII-1 SIRV1 gp29 (2 µM dimer), we incubated 4 µg *S. solfataricus* Csm complex (roughly 140 nM Csm carrying crRNA targeting A26 RNA target) with A26 RNA target (50 nM, 20 nM, 5 nM, 2 nM or 0.5 nM) in buffer containing 20 mM MES pH 6.0, 100 mM NaCl, 1 mM DTT and three units SUPERase-In Inhibitor supplemented with 2 mM MgCl<sub>2</sub> and 0.5 mM ATP at 70 °C for 60 min. We added 5'-end <sup>32</sup>P-labelled A1 RNA (5'-AGGGUA-UUAUUUGUUUUGUUUCUUAACUAUAAGCUAGUUCUGGAGA-3') and 0.5 µM dimer SsoCsx1 to the reaction at 60 min, and allowed the reaction to proceed for a further 60 min before quenching by adding phenol chloroform. We visualized A1 RNA cleavage by phosphor imaging after denaturing PAGE. A control reaction incubating SsoCsx1 with A1 RNA in the absence of cOA was carried out to determine SsoCsx1 background activity. We visualized cA<sub>4</sub> synthesis by Csm in response to A26 target RNA, and subsequent cA<sub>4</sub> degradation in the presence of Crn1 Sso2081 or AcrIII-1 SIRV1 gp29, by adding 5 nM α-<sup>32</sup>P-ATP with 0.5 mM ATP at the start of the reaction. Reactions were quenched at 60 min with phenol chloroform, and cA<sub>4</sub> degradation products were visualized by phosphor imaging following TLC. We also carried out a control reaction incubating Csm with ATP and α-<sup>32</sup>P-ATP in the absence of A26 target RNA, quenching the reaction after 60 min.

We determined the cA<sub>4</sub>-degradation capacity of AcrIII-1 SIRV1 gp29 and of the Crn1 enzyme Sso2081 by incubating 2 µM dimer of each enzyme with 500–0.5 µM unlabelled cA<sub>4</sub> (BIOLOG Life Science Institute, Bremen, Germany) in Csx1 buffer at 70 °C for 20 min before introducing SsoCsx1 (0.5 µM dimer) and <sup>32</sup>P-labelled A1 RNA (50 nM). The reaction was left to proceed for a further 60 min at 70 °C before quenching by adding phenol chloroform. Deproteinized products were separated by denaturing PAGE to visualize RNA degradation.

### Plasmid immunity from a reprogrammed type III system

Plasmids pCsm1-5\_ΔCsm6 (containing the type III Csm interference genes *cas10*, *csml3*, *csml4* and *csml5* from *M. tuberculosis* and *csml2* from *M. canettii*), pCRISPR\_TetR (containing *M. tuberculosis cas6* and a tetracycline-resistance-gene-targeting CRISPR array), pRAT-Target (tetracycline-resistance plus target plasmid) and *M. tuberculosis (Mtb) Csm6/Thioalkalivibrio sulfidiphilus (Tsu) Csx1* expression constructs have been described previously<sup>21</sup>. pRAT-Duet was constructed by

replacing the pUC19 *lacZα* gene of pRAT-Target with the multiple cloning sites (MCSs) of pACYCDuet-1 by restriction digest (5'-NcoI, 3'-XhoI). The viral ring nuclease (*duf1874*) gene from *Thermoanaerobacterium* phage THSA\_485A, tsac\_2833, was PCR-amplified from its pHisTEV expression construct and cloned into the 5'-NdeI, 3'-XhoI sites of MCS-2. The cOA-dependent nuclease genes (*mtb csm6*, *tsu csx1*) were cloned into the 5'-NcoI, 3'-SalI sites of MCS-1 by restriction digest from their respective expression constructs. Each nuclease was cloned with and without the viral ring nuclease; pRAT-Duet without insert and pRAT-Duet containing only the viral ring nuclease were used as controls. We carried out the plasmid transformation assay essentially as described<sup>21</sup>. *E. coli* C43 containing pCsm1-5\_ΔCsm6 and pCRISPR\_TetR were transformed by heat shock with 100 ng of pRAT-Duet target plasmid containing different combinations of cOA-dependent nuclease and viral ring nuclease. After outgrowth at 37 °C for 2 h, cells were collected and resuspended in 200 µl LB. A series of tenfold dilutions was applied onto LB agar containing 100 µg ml<sup>-1</sup> ampicillin and 50 µg ml<sup>-1</sup> spectinomycin to determine the cell density of the recipient cells and onto LB agar additionally containing 25 µg ml<sup>-1</sup> tetracycline, 0.2% (*w/v*) D-lactose and 0.2% (*w/v*) L-arabinose to determine the cell density of viable transformants. Plates were incubated at 37 °C for 16–18 h; further incubation was carried out at room temperature. Colonies were counted manually and corrected for dilution and volume to obtain colony-forming units (CFUs) per millilitre. Raw data for plasmid counts are available in Supplementary Data 3.

### Liquid chromatography/high-resolution mass spectrometry

We incubated AcrIII-1 SIRV1 gp29 (40 µM dimer) with 400 µM cA<sub>4</sub> in Csx1 buffer for 2 min at 70 °C, and carried out deproteinization by phenol-chloroform extraction followed by chloroform extraction. Liquid chromatography/high-resolution mass spectrometry (LC-HRMS) analysis was performed on a Thermo Scientific Velos Pro instrument equipped with HESI source and Dionex UltiMate 3000 chromatography system. Compounds were separated on a Kinetex EVO C18 column (2.6 µm, 2.1 × 50 mm; Phenomenex) using the following gradient of acetonitrile (B) against 20 mM ammonium bicarbonate (A): 0–2 min 2% B, 2–10 min 2–8% B, 10–11 min 8–95% B, 11–14 min 95% B, 14–15 min 95–2% B, 15–20 min 2% B, at a flow rate of 300 µl min<sup>-1</sup> and column temperature of 40 °C. UV data were recorded at 254 nm. Mass data were acquired on a Fourier transform mass analyser in negative-ion mode, with scan range *m/z* 150–1,500 at a resolution of 30,000. We set the source voltage to 3.5 kV, the capillary temperature to 350 °C, and the source heater temperature to 250 °C. Data were analysed using Xcalibur (Thermo Scientific).

### Phylogenetic analysis

AcrIII-1 homologues were collected by using gp29 (NP\_666617) of SIRV1 as a query and running two iterations ( $E = 1 \times 10^{-5}$ ) of PSI-BLAST<sup>39</sup> against the non-redundant protein database at the National Center for Biotechnology Information (NCBI). Sequences were aligned using PROMALS3D<sup>40</sup>. Redundant sequences (95% identity threshold) and sequences with a mutated active-site residue H47 were removed from the alignment. Poorly aligned (low information content) positions were removed using the *gt 0.2* function of Trimal<sup>41</sup>. The final alignment contained 124 positions. The maximum likelihood phylogenetic tree was constructed using PhyML<sup>42</sup> with automatic selection of the best-fit substitution model for a given alignment. The best model identified by PhyML was LG + G + I. We assessed branch support using aBayes implemented in PhyML, and visualized the tree using iTOL<sup>43</sup>.

### Crystallization

The AcrIII-1 H47A variant was concentrated to 10 mg ml<sup>-1</sup>, incubated at 293 K for 1 h with a 1.2 M excess of cA<sub>4</sub>, and centrifuged at 13,000 r.p.m. for 10 min before crystallization. Sitting drop vapour diffusion experiments were set up at the nanolitre scale using commercially available

# Article

and in-house crystallization screens and incubated at 293 K. Crystals appeared in various conditions, but those used for data collection grew from 40% 2-methyl-2,4-pentanediol, 5% polyethylene glycol 8000 and 0.1 M sodium cacodylate, pH 6.5. Crystals were harvested and transferred briefly into cryoprotectant containing mother liquor with 20% glycerol immediately before cryo-cooling in liquid nitrogen. We used the H47A variant to avoid cleavage of the cA<sub>4</sub> substrate during co-crystallization. The position of the active-site histidine was inferred from the structure of the apo-protein.

## Data collection and processing

X-ray data were collected from two crystals at 100 K, at a wavelength 0.9686 Å, on beamline I24 at the Diamond Light Source, to 1.49 Å and 1.60 Å resolution. Both data sets were automatically processed with Xia2<sup>44</sup>, using XDS and XSCALE<sup>45</sup>. The data were merged in Aimless<sup>46</sup> and the overall resolution truncated to 1.55 Å. The data were phased by molecular replacement using Phaser<sup>47</sup>, with a monomer from PDB file 2X4I stripped of water molecules as the search model. Model refinement of AcrIII-1 was achieved by iterative cycles of REFMAC5<sup>48</sup> in the CCP4 suite<sup>49</sup> and manual manipulation in COOT<sup>50</sup>. Electron density for cA<sub>4</sub> was clearly visible in the maximum likelihood/ $\sigma_A$ -weighted  $F_{\text{obs}} - F_{\text{calc}}$  electron-density map at 3 $\sigma$ . The coordinates for cA<sub>4</sub> were generated in ChemDraw (Perkin Elmer) and the library was generated using acedrg<sup>51</sup>, before fitting of the molecule in COOT. Model quality was monitored throughout using Molprobity<sup>52</sup> (score 1.13; centile 99). Ramachandran statistics were 98.5% favoured, 0% disallowed. Data and refinement statistics are shown in Extended Data Table 1.

## Sample size and randomization

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

## Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

## Data availability

The structural coordinates and data have been deposited in the Protein Data Bank (PDB) with deposition code 6SCF. The genome sequence of the SSeV virus has been submitted to GenBank with accession code MN53972. Raw data are available in the Supplementary Information for the plasmid immunity analysis presented in Fig. 1 and Extended Data Fig. 3, and the kinetic analysis presented in Fig. 2 and Extended Data Figs. 5, 6.

31. Held, N. L., Herrera, A. & Whitaker, R. J. Reassortment of CRISPR repeat-spacer loci in *Sulfolobus islandicus*. *Environ. Microbiol.* **15**, 3065–3076 (2013).
32. Zhang, C. & Whitaker, R. J. Microhomology-mediated high-throughput gene inactivation strategy for the hyperthermophilic crenarchaeon *Sulfolobus islandicus*. *Appl. Environ. Microbiol.* **84**, e02167-17 (2017).
33. Zhang, C., Cooper, T. E., Krause, D. J. & Whitaker, R. J. Augmenting the genetic toolbox for *Sulfolobus islandicus* with a stringent positive selectable marker for agmatine prototrophy. *Appl. Environ. Microbiol.* **79**, 5539–5549 (2013).

34. Deng, L., Zhu, H., Chen, Z., Liang, Y. X. & She, Q. Unmarked gene deletion and host-vector system for the hyperthermophilic crenarchaeon *Sulfolobus islandicus*. *Extremophiles* **13**, 735–746 (2009).
35. Rouillon, C., Athukoralage, J. S., Graham, S., Grischow, S. & White, M. F. Investigation of the cyclic oligoadenylate signaling pathway of type III CRISPR systems. *Methods Enzymol.* **616**, 191–218 (2019).
36. Linkert, M. et al. Metadata matters: access to image data in the real world. *J. Cell Biol.* **189**, 777–782 (2010).
37. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
38. Sternberg, S. H., Haurwitz, R. E. & Doudna, J. A. Mechanism of substrate selection by a highly specific CRISPR endoribonuclease. *RNA* **18**, 661–672 (2012).
39. Altschul, S. F. et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402 (1997).
40. Pei, J. & Grishin, N. V. PROMALS3D: multiple protein sequence alignment enhanced with evolutionary and three-dimensional structural information. *Methods Mol. Biol.* **1079**, 263–271 (2014).
41. Capella-Gutiérrez, S., Silla-Martínez, J. M. & Gabaldón, T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972–1973 (2009).
42. Guindon, S. et al. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* **59**, 307–321 (2010).
43. Letunic, I. & Bork, P. Interactive tree of life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.* **47** (W1), W256–W259 (2019).
44. Winter, G. xia2: an expert system for macromolecular crystallography data reduction. *J. Appl. Crystallogr.* **43**, 186–190 (2010).
45. Kabsch, W. Xds. *Acta Crystallogr. D* **66**, 125–132 (2010).
46. Evans, P. R. An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr. D* **67**, 282–292 (2011).
47. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
48. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D* **53**, 240–255 (1997).
49. Winn, M. D. et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr. D* **67**, 235–242 (2011).
50. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
51. Long, F. et al. AceDRG: a stereochemical description generator for ligands. *Acta Crystallogr. D* **73**, 112–122 (2017).
52. Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D* **66**, 12–21 (2010).
53. Gerlt, J. A. Genomic enzymology: web tools for leveraging protein family sequence-function space and genome context to discover novel functions. *Biochemistry* **56**, 4293–4308 (2017).

**Acknowledgements** This work was supported by grants from the Biotechnology and Biological Sciences Research Council (BB/S000313/1 to M.F.W. and BB/R008035/1 to T.M.G.) and by a NASA Exobiology and Evolutionary Biology grant (NNX14AK23G to R.J.W.). We thank J. Black and M. Alejandra-Bautista for isolating and characterizing the SSeV virus, and R. Wipfler and W. Zhu for technical assistance.

**Author contributions** J.S.A. designed experiments and carried out enzyme assays and analysis; S.A.M. carried out structural biology; C.Z. constructed the *S. islandicus* strains and performed virus infection assays; Sabine Grischow carried out plasmid transformation assays and mass spectrometry; Shirley Graham generated expression plasmids and purified proteins; M.K. contributed to the conception of the project and performed phylogenetic analysis; T.M.G., R.J.W. and M.F.W. oversaw the work, analysed the data and wrote the manuscript. All authors contributed to data analysis and writing.

**Competing interests** The University of St Andrews has filed a patent application (UK Patent Application 1902256.5, “Novel enzyme for phage therapy”; filed 19 February 2019), on which J.S.A. and M.F.W. are inventors. The other authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41586-019-1909-5>.

**Correspondence and requests for materials** should be addressed to T.M.G. or M.F.W.

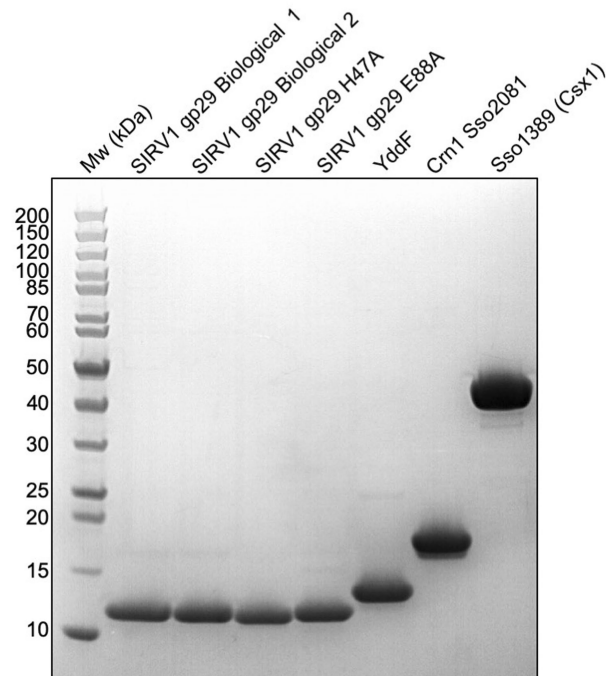
**Peer review information** Nature thanks Joseph Bondy-Denomy and John van der Oost for their contribution to the peer review of this work.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>.

**a**

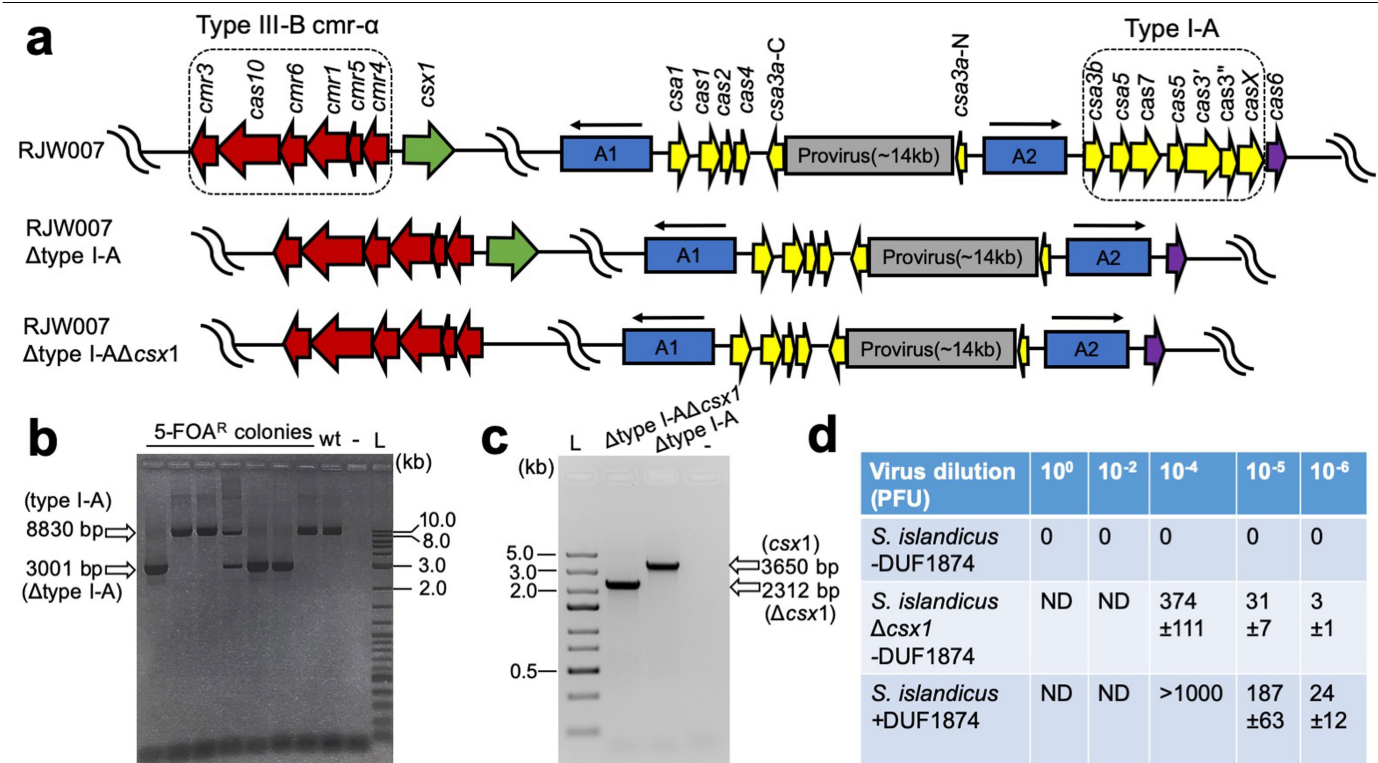
SIRV1_gp29	1	MNKVYLANAFSINM--LTKFPTKVVIDKIDRLEFCEN---IDNED	40
STIV_B116	1	MGKVFLTNAFSINM--LKEFPTTITIDKLDEEDFCLKLELRLEDGT	44
AFV3-109	1	MLYILNSAILPL--KPGEEYTVKAKEITIQEAKEL---VTKEQ	38
ARV1_gp13	1	MLYILNAQITP--FEGAQATFVERRIDVNEAKKI---VNSQP	37
SIFV_118	1	MLYILNSATLPL--KPGKEYVIHAKELTIEEAKEL---LENER	38
SMV4_113	1	MTVYLANAFSPSM--LNKLPSAVEFQRVDQKEFCEA---IHHG	38
ATV_gp06	1	MGVWSVVLNLTIVP--FRDERAKFEIERVSAEEAKKIQ--MHNSQ	45
Thermoanaerobacterium phage THSA-485A	1	MFIAANAFSLQM--LSQFPAHIDIEEVATSAVAKL-----D	33
Synechococcus phage S-CBWM1	1	MACCVVP--KGAPGLWSVVEISLEEVIQD---LEEGE	32
Fusobacterium phage Fnu1	1	MTIGILNTPILT---GEGTYKLSNITLEQAQKL---VNENE	35
Hydrogenobaculum phage 1	1	MLYVLNLSLIVPVDQNKQGYIVSLWKIDLETARKI---VREMP	40
ICEBs1 Yddf	1	MEIAFLNLSLVVT---SPGFYKAEKITLDELKHY---DGR	33
Crenothrix polyspora	1	MTLFIINAPILT---SYGDWRFEFEG-PLSIDKTRKL---LR-EG	35
-----N-----			
SIRV1_gp29	41	IINSIGHDSTIQLINSLCGTTFQKNRVEIKLEKEDKLYVVQI-SQRLEEG	89
STIV_B116	45	LINAI GHDS TINLVNTLCGTQLQKNRVEVKMNEGDEALIMI-SQRLEEG	93
AFV3-109	39	FTSAIGHQATAELLSSILGVNVPMNRVQIKVTHGDRILAFML-KQRLEEG	87
ARV1_gp13	38	FVSAVGHAAATAQLLSKLLDASIPTNRRTQVFLKPGDMALAIVL-KSRIPEG	86
SIFV_118	39	FISAVGHEATAKMLTNIFDVEIPMNRRIQIFLDDGDKLLSIIL-KTRLLEG	87
SMV4_113	39	VSNAI GHKGTIEFVNTLCNTNLQTNRVEIKAGINDVIYIIVL-GFRLEEG	87
ATV_gp06	46	FVSAIGHASANALSLLGVAVPVNRTEVFFNVGDEAIAIAML-KKRLAEG	94
Thermoanaerobacterium phage THSA-485A	34	LQSAIGHADTAVVLSGILGKDIESNRNVNQLQPGDSLIVAQLMGGRLEEG	83
Synechococcus phage S-CBWM1	33	FISTIGHPSSAHILETLTGFPFEACRRREADPRPGDEFYCFIL-NSRAPEG	81
Fusobacterium phage Fnu1	36	FISYIGHQATAEIIISILLGTEVPMNRGQFKQEVGQKAIIFKL-KSRLLEG	84
Hydrogenobaculum phage 1	41	FTSAVGHAEATAKVLSELLGVEISFNRIITVKMKEGDAGLHFVL-RTLPEEG	89
ICEBs1 Yddf	34	YKSFIGHKSTAQFLQKLLGIRIEQNRKTFRHMKYQKAI CFSL-YERYPEN	82
Crenothrix polyspora	36	FTSAIGHAASAEMLARLLAMDIPVNRITAITMEAGDRALILRL-LQRLPEEG	84
-----R-EG			
SIRV1_gp29	90	KILTLEEEILKLYESGKVQFFFEIIVD	114
STIV_B116	94	KVLSDEKEIKDMYRQGGKISFYEVW	116
AFV3-109	88	VVVKTTTEELEKI---GYELWLFEIQ	109
ARV1_gp13	87	VVLDE-QAIRNI---GFEIVVIERVS	108
SIFV_118	88	KVIKTVEELEKI---GYNIWLFEEVVTYEHNVKYE	118
SMV4_113	88	KVLSA-GEVQKAYDEGKVLLLKAIIGK	113
ATV_gp06	95	QVLRRTVQELEAV---GFDLYYIKRVQ	117
Thermoanaerobacterium phage THSA-485A	84	STTLP-AGF---SFKFFKVTVQA	102
Synechococcus phage S-CBWM1	82	KILDE-HEIYKI---GFSFRKMTYVLGKIPTAPD	111
Fusobacterium phage Fnu1	85	QILLTIQEIIEE---GYEFQLLERKN	107
Hydrogenobaculum phage 1	90	KVLSE-EELRQL---DFDLVLSRV	110
ICEBs1 Yddf	83	VLLTQ-RDLEKA---RYQFYLLTRL	104
Crenothrix polyspora	85	KVLNH-HEMMAT---PFELALLTKLK	106

**b**



**Extended Data Fig. 1 | Multiple sequence alignment of DUF1874-family members, and purity of DUF1874 and CRISPR ancillary enzymes used in biochemical assays. a.** This multiple sequence alignment includes the AcrIII-1 proteins from the archaeal viruses SIRV1, STIV, AFV3, ARV1, SIFV, SMV4 and ATV, the ICEBs1 protein Yddf from *B. subtilis*, the bacteriophage proteins from *Thermoanaerobacterium* phage THSA-485A, *Synechococcus* phage S-CBWM1, *Fusobacterium* phage Fnu1 and *Hydrogenobaculum* phage 1, and the Crn2

protein from *Crenothrix polyspora*. Conserved residues H47, R66, R85 and E88 are indicated by asterisks. Light and dark grey shading indicate regions of partial and strong sequence conservation, respectively. **b.** SDS-PAGE of SIRV1 gp29 (wild-type, H47A and E88A variants), Yddf, the Crn1 enzyme Sso2081, and the Csx1 enzyme Sso1389. The gel is representative of two or more biological replicates.

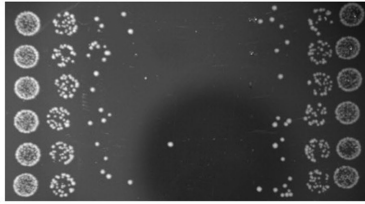


**Extended Data Fig. 2 | Construction of RJW007 Δtype I-A and RJW007 Δtype I-AΔ*csx1* mutant strains.** **a**, Genomic context of the CRISPR system in the genetic host (*S. islandicus* RJW007) and in mutant strains. A1 and A2 denote two different CRISPR arrays, the orientations of which are indicated with arrows. **b**, PCR verification of Δtype I-A mutants. A representative *Sulfolobus* transformant with integrated type I-A knockout plasmid was grown in dextrin-tryptone liquid medium, and the cell cultures were plated on dextrin-tryptone plates containing 5-fluoroorotic acid (5-FOA, 50 μg mg<sup>-1</sup>), uracil (20 μg ml<sup>-1</sup>), and agmatine (1 mg ml<sup>-1</sup>). Seven randomly selected 5-FOA-resistant (5-FOA<sup>R</sup>) colonies were screened using the primers that bind outside of the flanking homologous regions to confirm the type I-A deletion. A representative Δtype I-A mutant was further colony purified for subsequent experiments. The

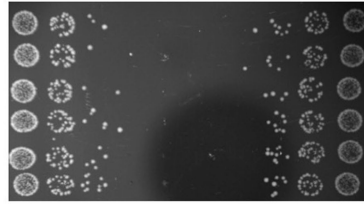
expected sizes of the PCR products amplified from the genomic DNA of the parental strain (referred to wild type, wt) and the Δtype I-A mutant are 8,830 base pairs (bp) and 3,001 bp, respectively. The minus symbol denotes a negative control (using water as the template for PCR). L, log-2 DNA ladder (NEB). Seven biological replicates were screened. **c**, PCR analysis of the RJW007 Δtype I-AΔ*csx1* mutant and its parental strain RJW007 Δtype I-A using primers that anneal to the outside of the flanking homologous regions of *csx1*, generating amplicons of 2,312 bp and 3,650 bp, respectively. Minus symbol, negative control (using water as the template for PCR). L, Gene Ruler Express DNA ladder (Thermo Scientific). The experiment carried out once. **d**, Plaque counts for the three strains tested (*n* = 3 biological replicates).

## Recipients

10<sup>-5</sup> 10<sup>-6</sup> 10<sup>-7</sup> 10<sup>-8</sup> 10<sup>-9</sup> 10<sup>-9</sup> 10<sup>-8</sup> 10<sup>-7</sup> 10<sup>-6</sup> 10<sup>-5</sup>



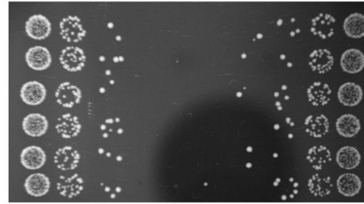
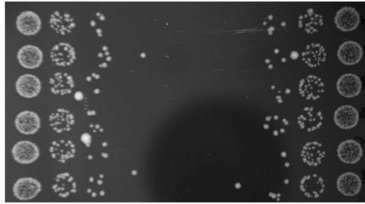
10<sup>-5</sup> 10<sup>-6</sup> 10<sup>-7</sup> 10<sup>-8</sup> 10<sup>-9</sup> 10<sup>-9</sup> 10<sup>-8</sup> 10<sup>-7</sup> 10<sup>-6</sup> 10<sup>-5</sup>



## 1 day

Nuclease: DUF1874:

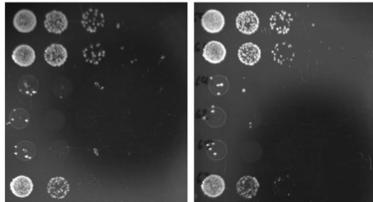
-	-
-	+
Csm6	-
Csm6	+
Csx1	-
Csx1	+



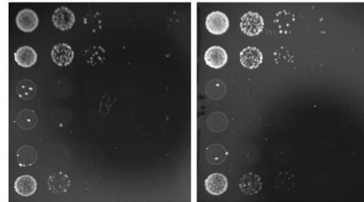
-	-
-	+
Csm6	-
Csm6	+
Csx1	-
Csx1	+

## Transformants (induced)

10<sup>-0</sup> 10<sup>-1</sup> 10<sup>-2</sup> 10<sup>-3</sup> 10<sup>-4</sup> 10<sup>-0</sup> 10<sup>-1</sup> 10<sup>-2</sup> 10<sup>-3</sup> 10<sup>-4</sup>



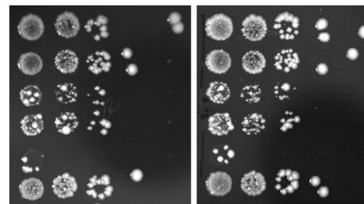
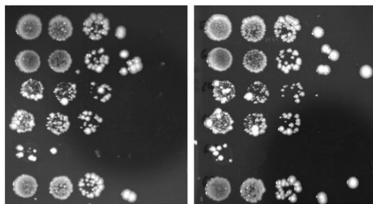
10<sup>-0</sup> 10<sup>-1</sup> 10<sup>-2</sup> 10<sup>-3</sup> 10<sup>-4</sup> 10<sup>-0</sup> 10<sup>-1</sup> 10<sup>-2</sup> 10<sup>-3</sup> 10<sup>-4</sup>



## 1 day

Nuclease: DUF1874:

-	-
-	+
Csm6	-
Csm6	+
Csx1	-
Csx1	+

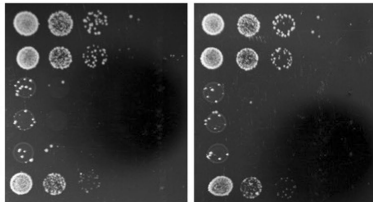


## 4 days

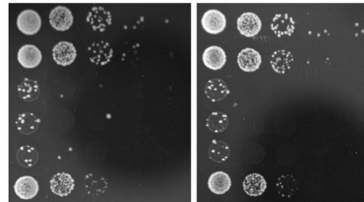
-	-
-	+
Csm6	-
Csm6	+
Csx1	-
Csx1	+

## Transformants (induced)

10<sup>-0</sup> 10<sup>-1</sup> 10<sup>-2</sup> 10<sup>-3</sup> 10<sup>-4</sup> 10<sup>-0</sup> 10<sup>-1</sup> 10<sup>-2</sup> 10<sup>-3</sup> 10<sup>-4</sup>



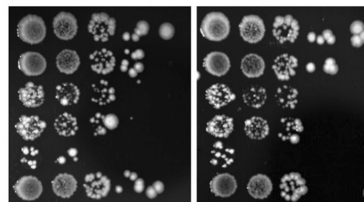
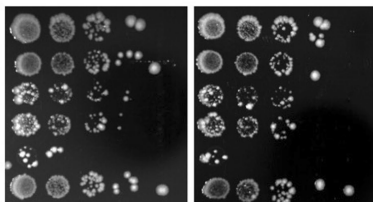
10<sup>-0</sup> 10<sup>-1</sup> 10<sup>-2</sup> 10<sup>-3</sup> 10<sup>-4</sup> 10<sup>-0</sup> 10<sup>-1</sup> 10<sup>-2</sup> 10<sup>-3</sup> 10<sup>-4</sup>



## 1 day

Nuclease: DUF1874:

-	-
-	+
Csm6	-
Csm6	+
Csx1	-
Csx1	+



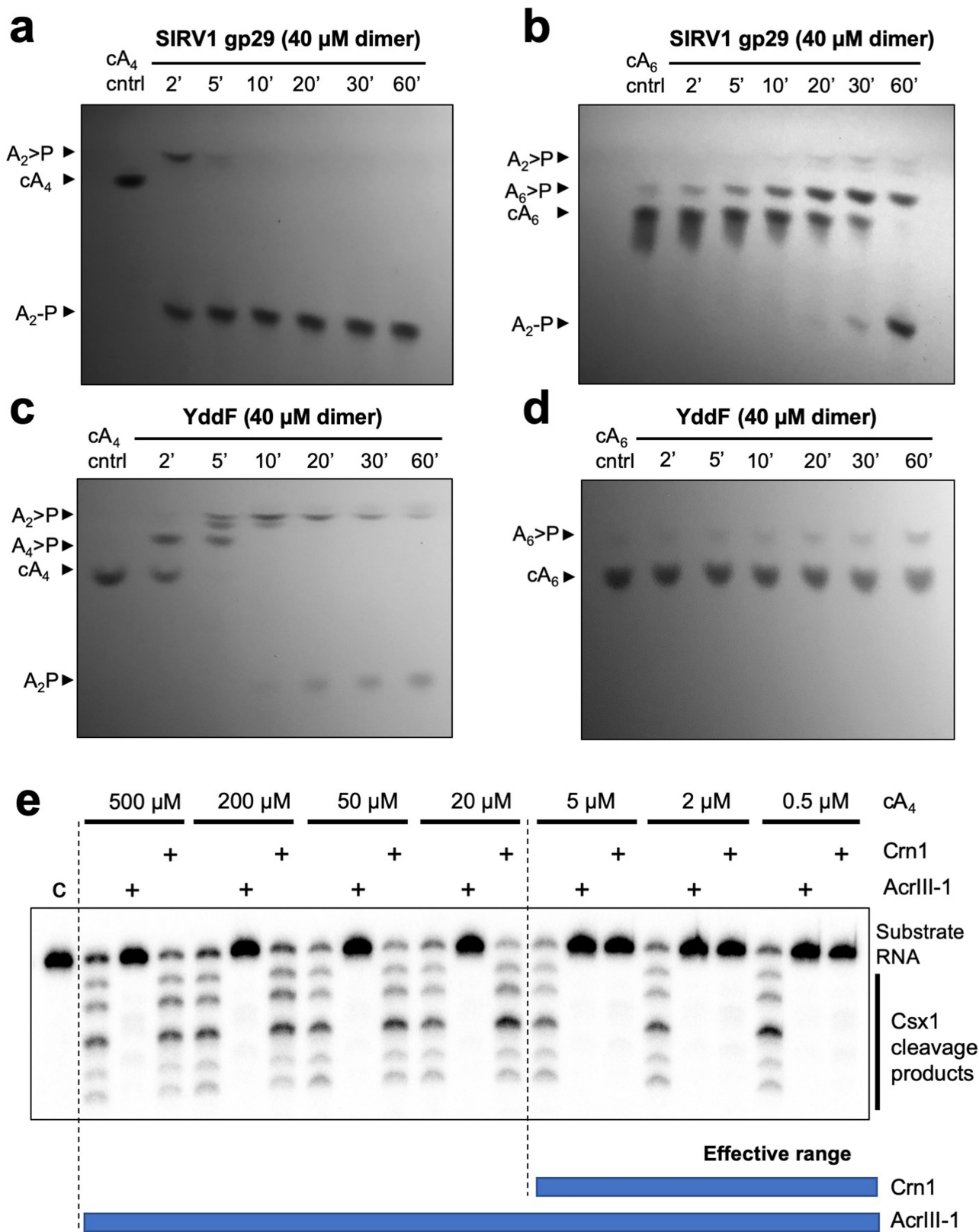
## 4 days

-	-
-	+
Csm6	-
Csm6	+
Csx1	-
Csx1	+

**Extended Data Fig. 3 | Effect of DUF1874 on plasmid immunity provided by a heterologously expressed *M. tuberculosis* type III-A CRISPR system, providing *CA<sub>4</sub>*- or *CA<sub>6</sub>*-mediated immunity.** Unprocessed images of sample

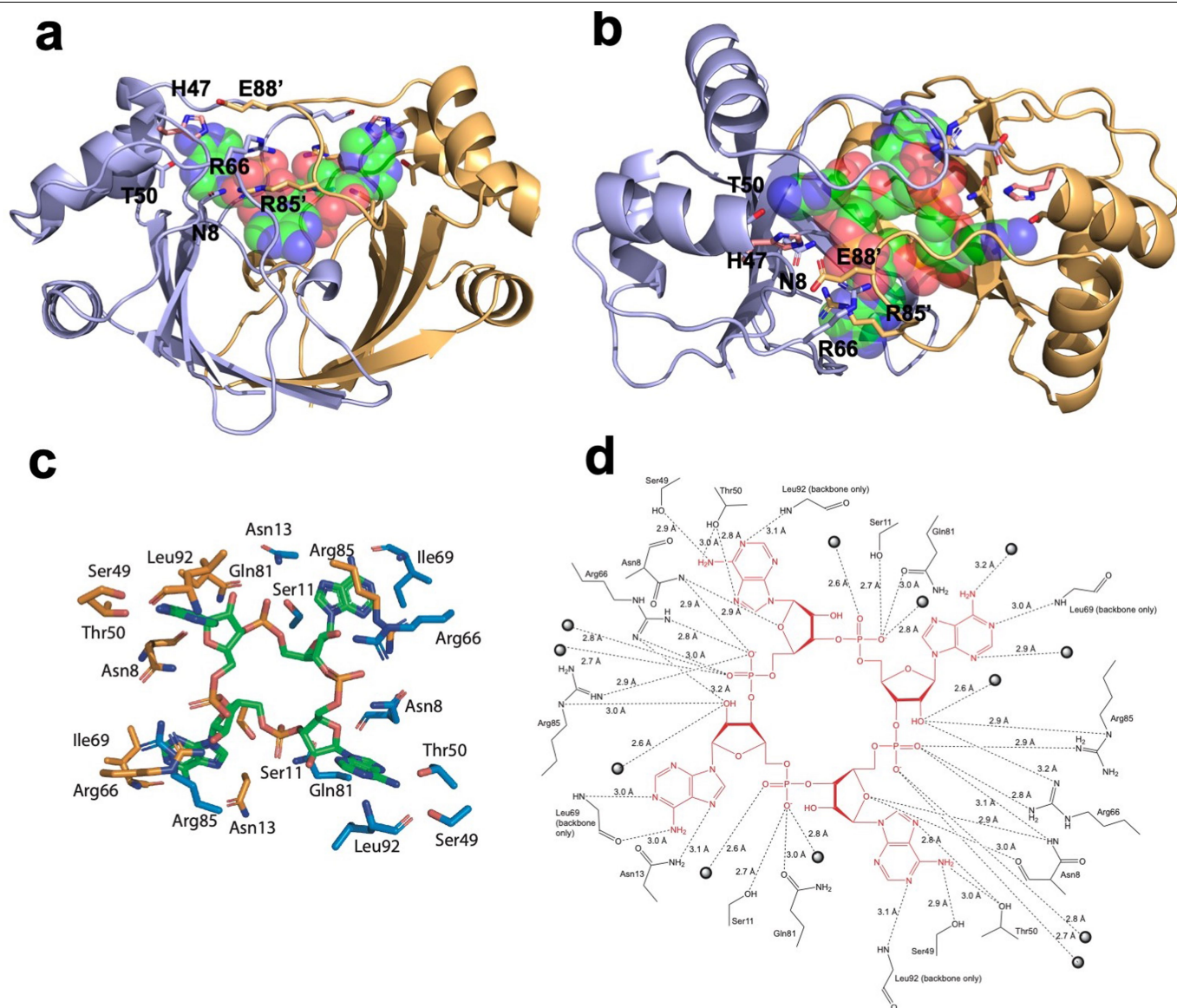
plates are shown for all replicates (two biological replicates with four technical replicates each; *n* = 8). Cell-culture dilutions are indicated above the plates.





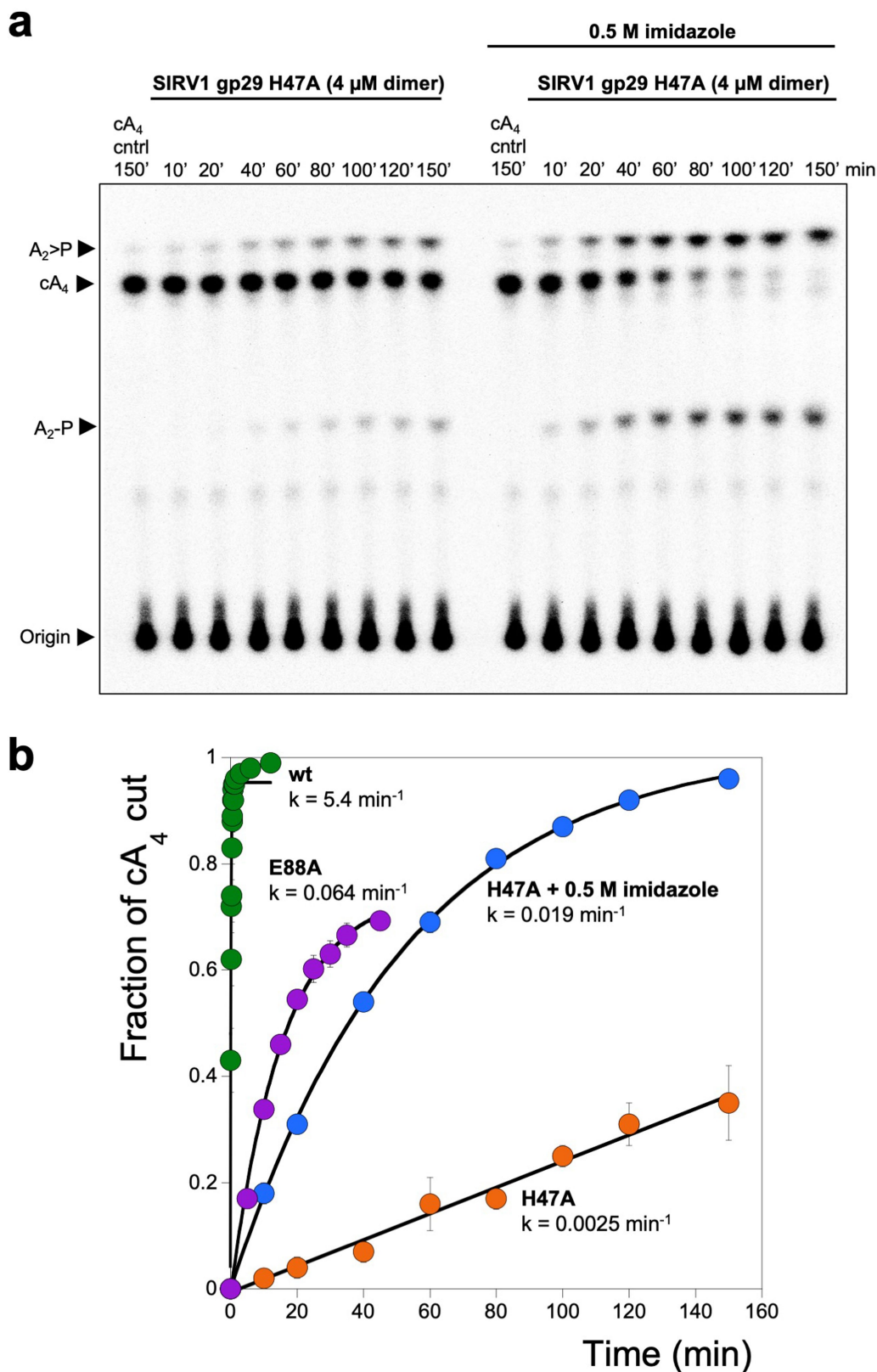
**Extended Data Fig. 4 | Substrate preference of the AcrIII-1 proteins SIRV1 gp29 and YddF, and effective range of  $cA_4$  degradation.** **a–d**, TLC images visualizing (under 254 nm UV light)  $cA_4$  and  $cA_6$  (450  $\mu$ M) degradation by SIRV1 gp29 (**a, b**) and YddF (**c, d**) over time (in minutes). Both AcrIII-1 enzymes display a clear preference for  $cA_4$  over  $cA_6$ . All TLC images are representative of three technical replicates. **e**, Denaturing PAGE showing activation of Csx1 (0.5  $\mu$ M

dimer) by the indicated amounts (500–0.5  $\mu$ M) of HPLC-purified  $cA_4$ , and its subsequent deactivation when either AcrIII-1 or Crn1 (2  $\mu$ M dimer) was present to degrade  $cA_4$ . The AcrIII-1 enzyme degraded 100-fold more  $cA_4$  than did Crn1. The control reaction (C) shows RNA incubated with Csx1 in the absence of  $cA_4$  ( $n = 3$  technical replicates). For gel source data, see Supplementary Fig. 1.



**Extended Data Fig. 5 | Structure of SIRV1 gp29 bound to cA<sub>4</sub>.** **a, b**, Orthogonal views of SIRV1 gp29 dimer in complex with cA<sub>4</sub>. The protein monomers are coloured purple and gold, with catalytic residue H47 from the apo structure shown in salmon. cA<sub>4</sub> is shown as a spacefill model, with green, blue, red and orange representing carbon, nitrogen, oxygen and phosphorus atoms,

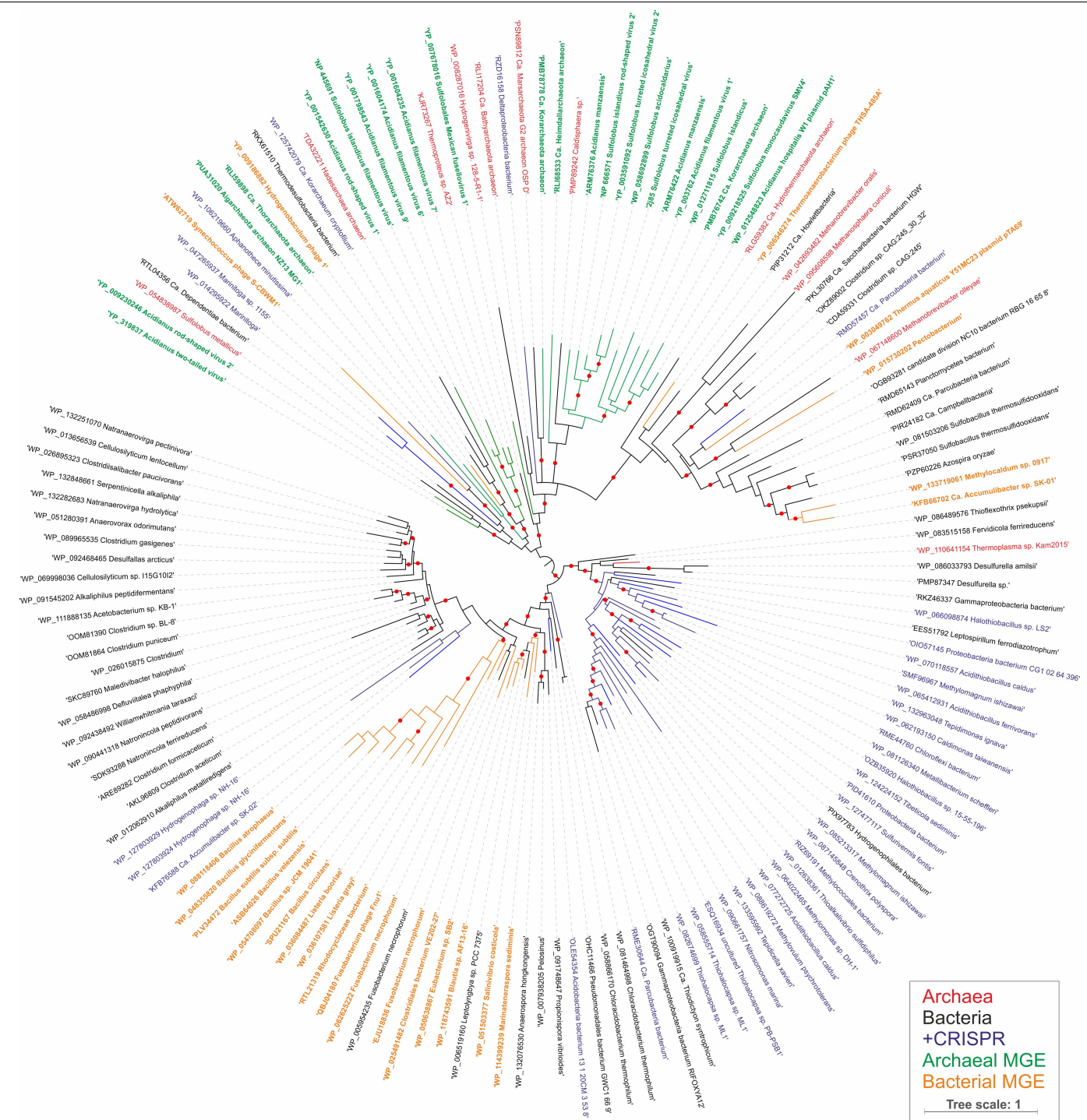
respectively. Conserved residues (Extended Data Fig. 1) in the AcrIII-1 family are indicated and discussed in the text. **c**, Interactions between each monomer of the SIRV1 dimer (orange and blue), with cA<sub>4</sub> shown in green. **d**, Diagram showing the interaction between SIRV1 gp29 and cA<sub>4</sub>. Dotted lines represent hydrogen bonds, with distances annotated. Spheres represent water molecules.



**Extended Data Fig. 6 | Single-turnover cA<sub>4</sub> cleavage by SIRV1 gp29 and variants, and chemical rescue with imidazole. a,** Phosphorimage of TLC visualizing cA<sub>4</sub> cleavage by SIRV1 gp29 H47A (4  $\mu$ M dimer, 50 °C) in the presence or absence of 500 mM imidazole, over time. The rate of cA<sub>4</sub> cleavage to generate A<sub>2</sub>>P and A<sub>2</sub>-P was calculated by quantifying densitometric signals from the phosphorimage ( $n = 3$  technical replicates). **b,** Plot comparing the

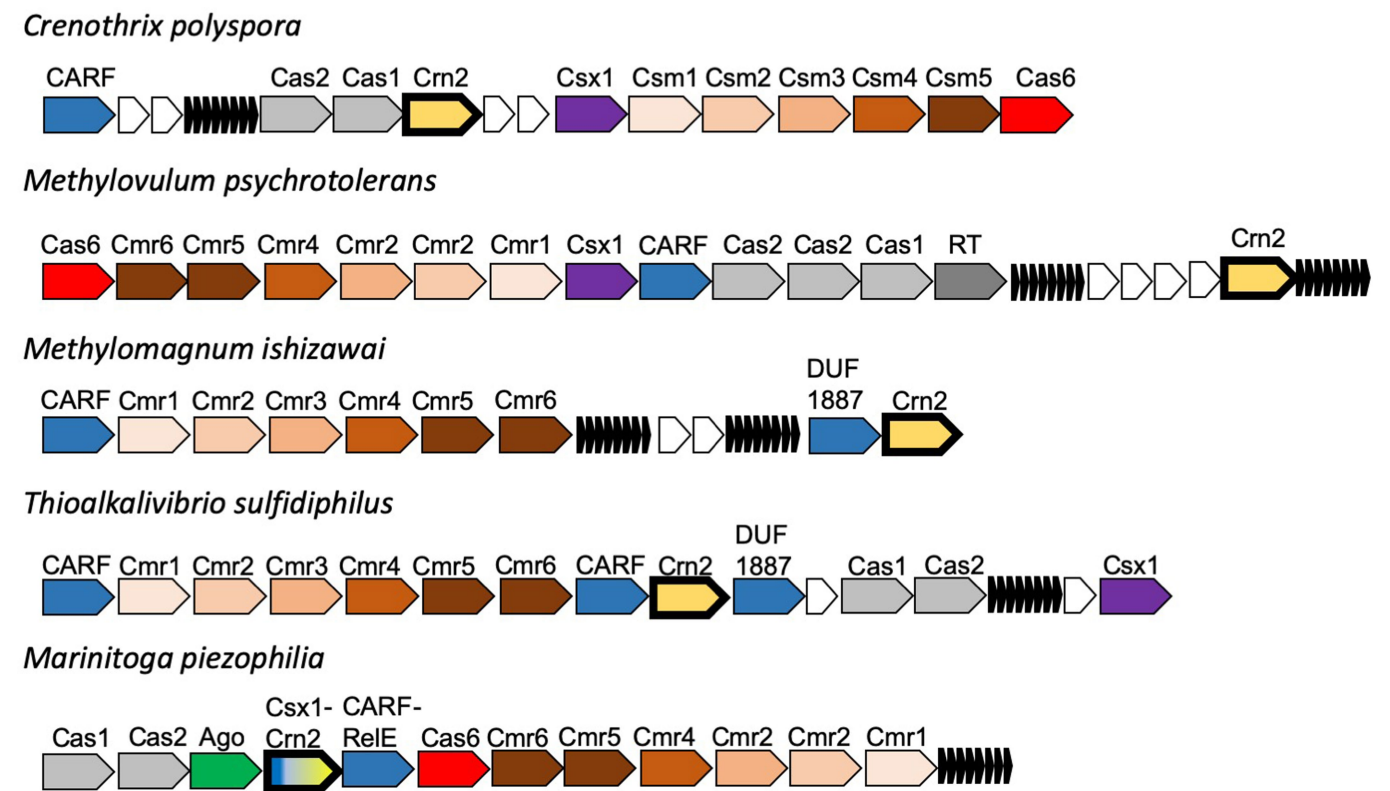
single-turnover rates of cA<sub>4</sub> by SIRV1 gp29, its E88A variant and its H47A variant, in the presence or absence of imidazole. Cleavage of cA<sub>4</sub> by the H47A variant can be partially restored when the reaction is supplemented with 500 mM imidazole. Data are mean and s.d. ( $n = 3$  technical replicates). For gel source data, see Supplementary Fig. 1.





**Extended Data Fig. 7 | Maximum likelihood phylogeny of AcrIII-1 homologues.** The maximum likelihood phylogenetic tree was constructed with automatic selection of the best-fit substitution model for a given alignment (LG + G + I). Red circles indicate 95–100% branch support, as assessed using aBayes implemented in PhyML. The scale bar represents the

number of substitutions per site. Branches and labels are colour coded: red, archaea; black, bacteria; blue, bacteria and archaea in which AcrIII-1 homologues are associated with CRISPR loci; green, archaeal viruses and plasmids; orange, bacteriophages.



**Extended Data Fig. 8 | Genomic context of crn2 genes in selected bacteria.** Type III CRISPR loci in the bacterial species *Crenothrix polyspora*, *Methylovulum psychrotolerans*, *Methylomagnum ishizawai*, *Thioalkalivibrio sulfidiphilus* and *Marinitoga piezophila* are shown, with genes labelled and colour coded. The crn2 gene is shown in pale yellow with a bold outline;

CRISPRs are indicated by small black arrowheads; and unrelated/hypothetical genes are shown as small white arrows. The sizes and orientations of genes are not reflected. Ago, Argonaute; CARF, CRISPR-associated Rossman fold; CARF-RelE, CARF domain fused to the RelE toxin; DUF1887, predicted CARF nuclease; RT, reverse transcriptase.





**Extended Data Fig. 9 | CRISPR-associated AcrIII-1 homologues.** Genomic neighbourhoods were analysed using the enzyme function initiative–genome neighbourhood tool (EFI-GNT) against the Pfam profile database<sup>53</sup>. Gene annotations are colour coded according to the key at the right.

Extended Data Table 1 | Data collection and refinement statistics for AcrIII-1 in complex with cA<sub>4</sub>

AcrIII-1 with cA <sub>4</sub>	
Data collection	
Space group	<i>P</i> 1
Cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	49.8, 51.7, 85.6
α, β, γ (°)	80.2, 89.7, 83.4
Resolution (Å)	50.63-1.55 (1.58-1.55) *
<i>R</i> <sub>sym</sub> or <i>R</i> <sub>merge</sub>	0.12 (0.36)
<i>I</i> / σ <i>I</i>	12.3 (1.7)
Completeness (%)	98.6 (92.4)
Redundancy	2.9 (1.8)
Refinement	
Resolution (Å)	84.26-1.55
No. reflections	113882
<i>R</i> <sub>work</sub> / <i>R</i> <sub>free</sub>	0.20 / 0.25
No. atoms	
Protein	7,365
Ligand/ion	352
Water	595
<i>B</i> -factors	
Protein	20.2
Ligand/ion	13.3
Water	30.7
R.m.s. deviations	
Bond lengths (Å)	0.012
Bond angles (°)	1.64

\*Values in parentheses are for the highest-resolution shell. *R*, residual factor; *I*, intensity.

## Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- |                                     |  |
|-------------------------------------|--|
| n/a                                 | Confirmed  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided<br><i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A description of all covariates tested  |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons   |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For null hypothesis testing, the test statistic (e.g. $F$ , $t$ , $r$ ) with confidence intervals, effect sizes, degrees of freedom and $P$ value noted<br><i>Give <math>P</math> values as exact values whenever suitable.</i>                                       |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings  |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes  |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Estimates of effect sizes (e.g. Cohen's $d$ , Pearson's $r$ ), indicating how they were calculated  |

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

### Software and code

Policy information about [availability of computer code](#)

#### Data collection

Provide a description of all commercial, open source and custom code used to collect the data in this study, specifying the version used OR state that no software was used.

#### Data analysis

Provide a description of all commercial, open source and custom code used to analyse the data in this study, specifying the version used OR state that no software was used.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

### Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The structural coordinates and data have been deposited in the Protein Data Bank with deposition code 6SCF.

Raw data is available for the plasmid immunity analysis presented in figure 1.

Raw data is available for the kinetic analysis presented in figure 2 and extended data figure 6.

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences      ☐ Behavioural & social sciences      ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	this is not relevant to the study
Data exclusions	no data were excluded
Replication	biological and technical replicates of kinetic assays were carried out
Randomization	this is not relevant to the study
Blinding	this is not relevant to the study

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data

### Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

# H2A.Z facilitates licensing and activation of early replication origins

<https://doi.org/10.1038/s41586-019-1877-9>

Received: 5 December 2018

Accepted: 31 October 2019

Published online: 25 December 2019

There are amendments to this paper

Haizhen Long<sup>1,2,10</sup>, Liwei Zhang<sup>1,10</sup>, Mengjie Lv<sup>3,10</sup>, Zengqi Wen<sup>1,2,10</sup>, Wenhao Zhang<sup>4</sup>, Xiulan Chen<sup>2,5</sup>, Peitao Zhang<sup>6</sup>, Tongqing Li<sup>7</sup>, Luyuan Chang<sup>1,2</sup>, Caiwei Jin<sup>2,3</sup>, Guozhao Wu<sup>1,2</sup>, Xi Wang<sup>8</sup>, Fuquan Yang<sup>2,5</sup>, Jianfeng Pei<sup>7</sup>, Ping Chen<sup>1</sup>, Raphael Margueron<sup>9</sup>, Haiteng Deng<sup>4</sup>, Mingzhao Zhu<sup>2,3\*</sup> & Guohong Li<sup>1,2\*</sup>

DNA replication is a tightly regulated process that ensures the precise duplication of the genome during the cell cycle<sup>1</sup>. In eukaryotes, the licensing and activation of replication origins are regulated by both DNA sequence and chromatin features<sup>2</sup>. However, the chromatin-based regulatory mechanisms remain largely uncharacterized. Here we show that, in HeLa cells, nucleosomes containing the histone variant H2A.Z are enriched with histone H4 that is dimethylated on its lysine 20 residue (H4K20me2) and with bound origin-recognition complex (ORC). In vitro studies show that H2A.Z-containing nucleosomes bind directly to the histone lysine methyltransferase enzyme SUV420H1, promoting H4K20me2 deposition, which is in turn required for ORC1 binding. Genome-wide studies show that signals from H4K20me2, ORC1 and nascent DNA strands co-localize with H2A.Z, and that depletion of H2A.Z results in decreased H4K20me2, ORC1 and nascent-strand signals throughout the genome. H2A.Z-regulated replication origins have a higher firing efficiency and early replication timing compared with other origins. Our results suggest that the histone variant H2A.Z epigenetically regulates the licensing and activation of early replication origins and maintains replication timing through the SUV420H1–H4K20me2–ORC1 axis.

In eukaryotes, DNA-replication origins are first licensed in G1 phase by the pre-replication complex<sup>3</sup>; the licensed origins are then selectively activated during S phase<sup>4</sup>. In budding yeast, the ORC recognizes autonomously replicating sequences (ARSs) to achieve origin licensing<sup>4</sup>. In metazoans, which lack ARSs, replication origins are determined by both DNA sequence and chromatin-associated factors<sup>2</sup>. Of these chromatin features, it has been reported<sup>5</sup> that the histone modification H4K20me2 is recognized by ORC1. Given the broad distribution of H4K20me2 across the genome<sup>6</sup>, however, other factors must be involved to precisely define the function of H4K20me2 in DNA replication. Genome-wide studies have shown that the histone variant H2A.Z is also enriched at replication origins<sup>7,8</sup>. However, whether the enrichment of H2A.Z has a functional role during DNA replication remained unclear.

## H4K20me2 and ORC1 recruitment by H2A.Z

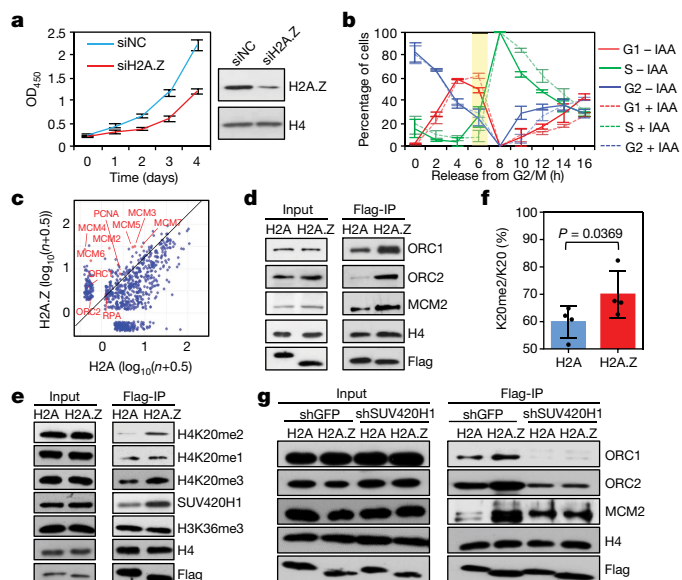
We first found that knocking down the *H2AFZ* genes in HeLa cells results in cell growth defects (Fig. 1a), but not in apoptosis or senescence (Extended Data Fig. 1a, b). Further analysis of *H2AFZ*-knockdown cells revealed a defect in incorporation of the replication marker bromodeoxyuridine (BrdU), along with a decreased proportion of S-phase

and an increased proportion of G1-phase cells (Extended Data Fig. 1c), indicating that the cells are arrested at the G1/S boundary. Next, we synchronized cells at G2/M phase and induced H2A.Z degradation using a knock-in auxin-inducible degron (AID) tag<sup>9</sup>. Six hours after release from G2/M arrest, H2A.Z-depleted cells (with depletion having been triggered with the auxin indole-3-acetate, or IAA; Extended Data Fig. 1d) showed a lower proportion of S-phase cells (dashed green line) and a higher proportion of G1-phase cells (dashed red line) (Fig. 1b). This finding supports the idea that cells become arrested at the G1/S boundary after H2A.Z depletion.

Although we identified some genes that are differentially expressed after *H2AFZ* knockdown (Extended Data Fig. 1e), we did not find any enriched terms relating to cell-cycle regulation. Reverse transcription–polymerase chain reaction (RT–PCR) analysis of genes involved in S-phase cell-cycle control (including *CCNE1* and *CDK2*) and DNA replication (including *ORC1*, *ORC2*, the DNA helicase *MCM2*, *CDC6*, *PCNA* and *RPA1*) revealed that *H2AFZ* knockdown did not change the expression of these genes by very much (Extended Data Fig. 1f). Next, using mass spectrometry, we found that the subunits of the pre-replication complex, including ORC1, ORC2 and MCM2–7, were enriched on H2A.Z nucleosomes compared with H2A nucleosomes (Fig. 1c and

<sup>1</sup>National Laboratory of Biomacromolecules, CAS Center for Excellence in Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing, China. <sup>2</sup>University of Chinese Academy of Sciences, Beijing, China. <sup>3</sup>Key Laboratory of Infection and Immunity, Institute of Biophysics, Chinese Academy of Sciences, Beijing, China. <sup>4</sup>Ministry of Education (MOE) Key Laboratory of Bioinformatics, School of Life Sciences, Tsinghua University, Beijing, China. <sup>5</sup>Laboratory of Proteomics, Institute of Biophysics, Chinese Academy of Sciences, Beijing, China. <sup>6</sup>Department of Cell Biology, Tianjin Medical University, Tianjin, China. <sup>7</sup>Center for Quantitative Biology, Academy for Advanced Interdisciplinary Studies, Peking University, Beijing, China. <sup>8</sup>Department of Immunology, School of Basic Medical Sciences, Capital Medical University, Beijing, China. <sup>9</sup>Institut Curie, PSL Research University, Paris, France. <sup>10</sup>These authors contributed equally: Haizhen Long, Liwei Zhang, Mengjie Lv, Zengqi Wen. \*e-mail: zhumz@sun5.ibp.ac.cn; liguohong@sun5.ibp.ac.cn





**Fig. 1 | H2A.Z interacts with the pre-replication complex.** **a**, Analysis of cell proliferation (left) and western blots (right) for HeLa cells transfected with negative-control short interfering (si) RNA (siNC) or siRNA that targets H2A.Z (siH2A.Z). OD<sub>450</sub>, optical density at 450 nm (an indicator of cell density). **b**, FACS analysis of the cell-cycle progression of control cells (–IAA; solid lines) and IAA-treated cells (+IAA; dash lines). The yellow shading shows that the cells were arrested at G1 phase and delayed proceeding into S phase after IAA-induced H2A.Z depletion. **c**, The total number of peptides identified from the immunoprecipitation of Flag-tagged H2A or H2A.Z nucleosomes from three independent experiments, plotted as log<sub>10</sub>(*n* + 0.5) with jitter. The diagonal line represents the threshold of twofold enrichment on H2A.Z nucleosomes. **d**, **e**, Western blot analysis of ORC1, ORC2 and MCM2 (**d**) or histone modifications and SUV420H1 (**e**) from immunoprecipitation of Flag-tagged H2A or H2A.Z mononucleosomes. **f**, Mass-spectrum analysis of the H4K20me2 modification from the samples in Fig. 1e. The y-axis shows the percentage of H4K20 peptides that are modified as H4K20me2. **g**, Western blot analysis of ORC1, ORC2 and MCM2 from immunoprecipitation of Flag-tagged H2A or H2A.Z mononucleosomes from cells stably expressing a control short hairpin (sh) RNA (shGFP) or shRNA against SUV420H1 (shSUV420H1). GFP, green fluorescent protein. Data in **a**, **b** are mean ± s.d.; **a**, *n* = 6 technical replicates; **b**, *n* = 3 biological replicates. Data in **f** are mean ± s.e.m.; *n* = 4 biological replicates, two-tailed, paired *t*-test. Western blots in **a**, **d**, **e**, **g** were independently repeated three times with similar results, and H4 was used as a loading control and sample processing control. For gel source data, see Supplementary Fig. 1. For the FACS gating strategy, see Supplementary Fig. 3.

Supplementary Table 1). We confirmed the enrichment of ORC1, ORC2 and MCM2 by western blotting (Fig. 1d).

As ORC1 might stabilize the binding of other ORC subunits at origins during G1 phase<sup>10</sup>, we tested the interaction between H2A.Z and ORC1 using the LacO/LacI targeting system. However, we did not find a direct interaction (Extended Data Fig. 1g), suggesting that a bridge is required to recruit ORC1 onto H2A.Z nucleosomes. The bromo adjacent homology (BAH) domain of ORC1 has been reported to specifically recognize the H4K20me2 peptide<sup>5</sup>, suggesting that ORC1 may be recruited onto H2A.Z nucleosomes through H4K20me2. Indeed, we found that both H4K20me2 and SUV420H1 were enriched on H2A.Z nucleosomes, in both unsynchronized (Fig. 1e) and G1-synchronized (Extended Data Fig. 1h) cells. Mass-spectrometry analysis of H2A and H2A.Z mononucleosomes showed that whereas H4K20me2 is abundant on both H2A and H2A.Z nucleosomes, it is relatively more enriched on the H2A.Z variant (Fig. 1f and Extended Data Fig. 2a, b). In addition, knockdown of *SUV420H1* alone or of both *SUV420H1* and *SUV420H2* (which encode the two enzymes<sup>11</sup> that catalyse methylation of H4K20me2) abolished the enrichment of ORC1, ORC2 and MCM2 on H2A.Z nucleosomes (Fig. 1g

and Extended Data Fig. 2c–e). These results suggest that H2A.Z recruits ORC1 in an H4K20me2-dependent manner. However, knockdown of *SUV420H2* alone had little effect on the binding of ORC1 onto H2A.Z nucleosomes (Extended Data Fig. 2d, e).

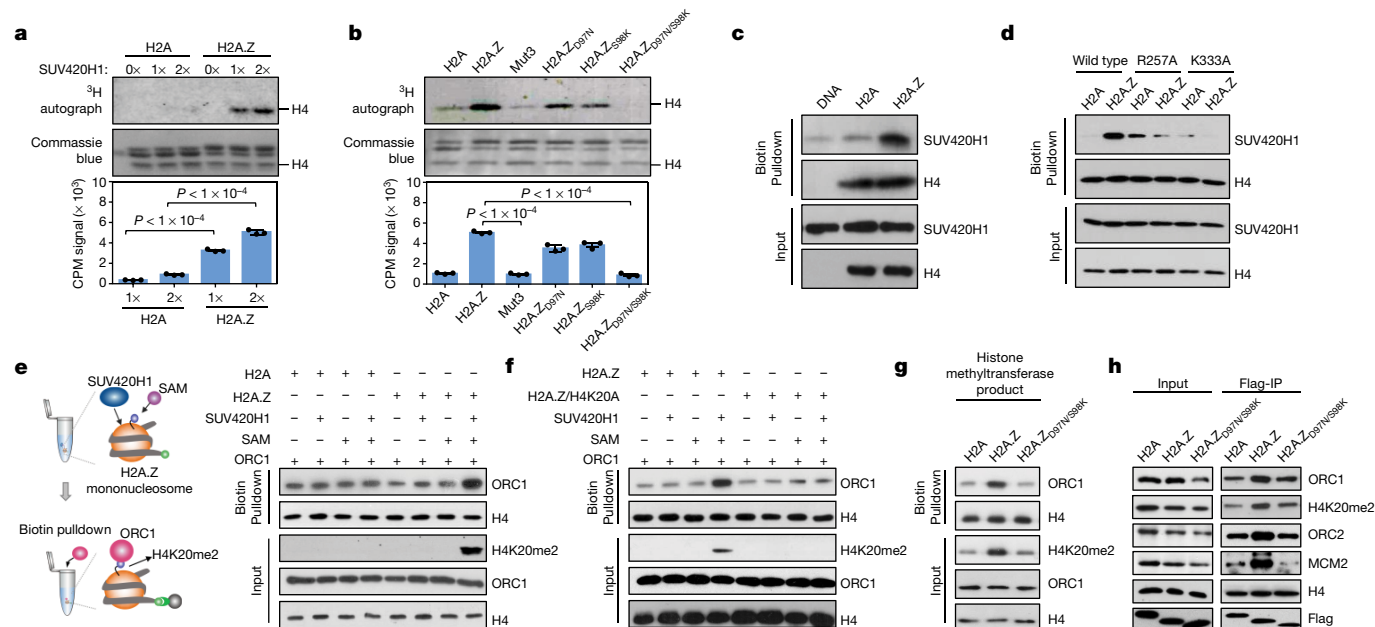
## H2A.Z recruits SUV420H1 to deposit H4K20me2

Methyltransferase assays showed that, compared with H2A mononucleosomes, H2A.Z mononucleosomes greatly enhanced the histone-methylation activity of recombinant human SUV420H1 (Fig. 2a and Extended Data Fig. 3a). Mass-spectrometry analysis of the modification products from the histone-methyltransferase reactions revealed that the main product of SUV420H1 activity on H2A.Z nucleosomes is H4K20me2, and that this is much more common on H2A.Z than on H2A nucleosomes (Extended Data Fig. 3b). We validated this result by western blotting (Extended Data Fig. 3c). It has been reported that SUV420H1 produces H4K20me2 from H4K20me1 *in vivo*<sup>12</sup>. In line with this, we found that H2A.Z also promoted the activity of SUV420H1 on nucleosomes containing H4K<sub>c</sub>20me1 (with K<sub>c</sub> being a lysine replaced by cystine for chemical modification; Extended Data Fig. 3d).

Next, we generated four chimaeric mutants of H2A.Z, containing regions that had been replaced by the corresponding regions of H2A (Extended Data Fig. 3e). When the acidic patch of H2A.Z was substituted with the corresponding domain of H2A, the activity of SUV420H1 was markedly reduced (Extended Data Fig. 3e). The residues D97 and S98 in the acidic patch have been reported<sup>13,14</sup> to be critical for the structural and biological functions of H2A.Z. Our results show that SUV420H1 has very low methylation activity on H2A.Z<sub>D97N/S98K</sub> mutant nucleosomes (Fig. 2b). Using mononucleosome pulldown assays, we found that SUV420H1 binds more strongly to H2A.Z nucleosomes than to H2A nucleosomes (Fig. 2c). Moreover, mutation of D97 and S98 to N97 and K98 in H2A.Z impaired the binding of SUV420H1 to nucleosomes (Extended Data Fig. 3f). We confirmed the binding of SUV420H1 to H2A.Z nucleosomes *in vivo* using the LacO/LacI targeting system (Extended Data Fig. 3g, h). Next, we simulated the binding between SUV420H1 and H2A.Z mononucleosomes using structural data for SUV420H1 (ref. 15) and the H2A.Z nucleosome<sup>16</sup>. We found that the R257 and K333 residues of SUV420H1 are important for the interaction with H2A.Z nucleosomes (Extended Data Fig. 4a, b). Indeed, both SUV420H1<sub>R257A</sub> and SUV420H1<sub>K333A</sub> mutants could not bind H2A.Z nucleosomes as efficiently as wild-type SUV420H1 (Fig. 2d). The methylation activity of these two mutants on H2A.Z nucleosomes was also reduced (Extended Data Fig. 4c). Together, these data show that residues D97 and S98 of H2A.Z, and R257 and K333 of SUV420H1, are essential for the binding and enhancing activity of SUV420H1.

A pulldown assay of biotinylated mononucleosomes showed that binding of ORC1 to H2A.Z nucleosomes is substantially enhanced by SUV420H1-catalysed H4K20me2 (Fig. 2e). Moreover, ORC1 binds weakly to the histone-methyltransferase products of H2A.Z/H4<sub>K20A</sub> nucleosomes (Fig. 2f) or H2A.Z<sub>D97N/S98K</sub> nucleosomes (Fig. 2g). We confirmed that the enrichment of H4K20me2, ORC1, ORC2 and MCM2, on H2A.Z nucleosomes is greatly impeded by H2A.Z<sub>D97N/S98K</sub> mutation *in vivo* (Fig. 2h). Using H2A.Z nucleosomes containing H4K<sub>c</sub>20me2 as substrates (Extended Data Fig. 4d), we found that mutations in the BAH domain of ORC1, which abolish the interaction between ORC1 and H4K20me2 (ref. 5), greatly impaired the interaction between ORC1 and H2A.Z nucleosomes containing H4K<sub>c</sub>20me2 (Extended Data Fig. 4e). These results support the idea that the binding of ORC1 to H2A.Z nucleosomes depends on the BAH domain of ORC1 and on SUV420H1-catalysed H4K20me2.

To test whether the density of H4K20me2 has an effect on ORC1 binding, we first assembled H2A and H2A.Z mononucleosomes containing zero (unmodified H4), one (heterotypic, 50%) or two (homotypic, 100%) H4 histones with the H4K<sub>c</sub>20me2 modification. As the H4K<sub>c</sub>20me2 density increased, ORC1 bound more strongly to H2A and



**Fig. 2 | H2A.Z binds SUV420H1 to promote H4K20me2 deposition and thereby recruit ORC1.** **a, b**, Upper panels,  $^3\text{H}$  autograph showing methyltransferase activity of SUV420H1 on wild-type H2A or H2A.Z mononucleosomes (**a**) and on mononucleosomes containing H2A.Z point mutations (**b**). Bottom, quantitative analysis of the  $^3\text{H}$  signal by liquid scintillation. Mut3 is an H2A.Z mutant whose acid patch is replaced with the corresponding region of H2A (Extended Data Fig. 3e). CPM, counts per minute. **c**, Western blot analysis of SUV420H1 from biotin pull-down samples. **d**, Western blot analysis of wild-type and mutant SUV420H1 from mononucleosome biotin pull-down samples. **e**, Left, diagram showing the histone methyltransferase activity of SUV420H1 on nucleosomes (top) and the subsequent binding of ORC1 to the products of SUV420H1's methyltransferase activity, detected by biotin pull-down. SAM, S-adenosyl methionine. Right, western blot analysis of

the binding of ORC1 to SUV420H1 products. **f**, Western blot analysis of the binding of ORC1 to SUV420H1's histone methyltransferase products, using H2A.Z mononucleosomes containing wild-type H4 or mutant H4<sub>K20A</sub> as substrates. **g**, Western blot analysis of the binding of ORC1 to SUV420H1 products, using wild-type H2A.Z or H2A.Z<sub>D97N/S98K</sub> mononucleosomes as substrates. **h**, Western blots showing the distribution of ORC1, ORC2, MCM2 and H4K20me2 on H2A, H2A.Z and H2A.Z<sub>D97N/S98K</sub> mononucleosomes.  $^3\text{H}$  quantification data in **a, b** are mean  $\pm$  s.e.m.;  $n = 3$  biological replicates; two-tailed, unpaired  $t$ -test. The  $^3\text{H}$  autograph experiments in **a, b** and western blots in **c–h** were independently repeated three times with similar results. H4 was used as a loading control and sample processing control. For gel source data, see Supplementary Fig. 1.

H2A.Z nucleosomes (Extended Data Fig. 4f). Next, we assembled H2A.Z polynucleosomes with an increasing ratio of an H4K<sub>20</sub>me2 octamer. We found that ORC1 binding was weak when the density of H4K<sub>20</sub>me2 was 25%, increasing gradually with increasing density of H4K<sub>20</sub>me2 (from 50% to 100%) (Extended Data Fig. 4g). These results suggest that ORC1 binds to chromatin in an H4K20me2-dosage-dependent manner. Thus, H2A.Z binds SUV420H1 directly to promote H4K20me2 deposition on H2A.Z nucleosomes, and the enhanced H4K20me2 deposition is essential for recruiting ORC1 to H2A.Z nucleosomes.

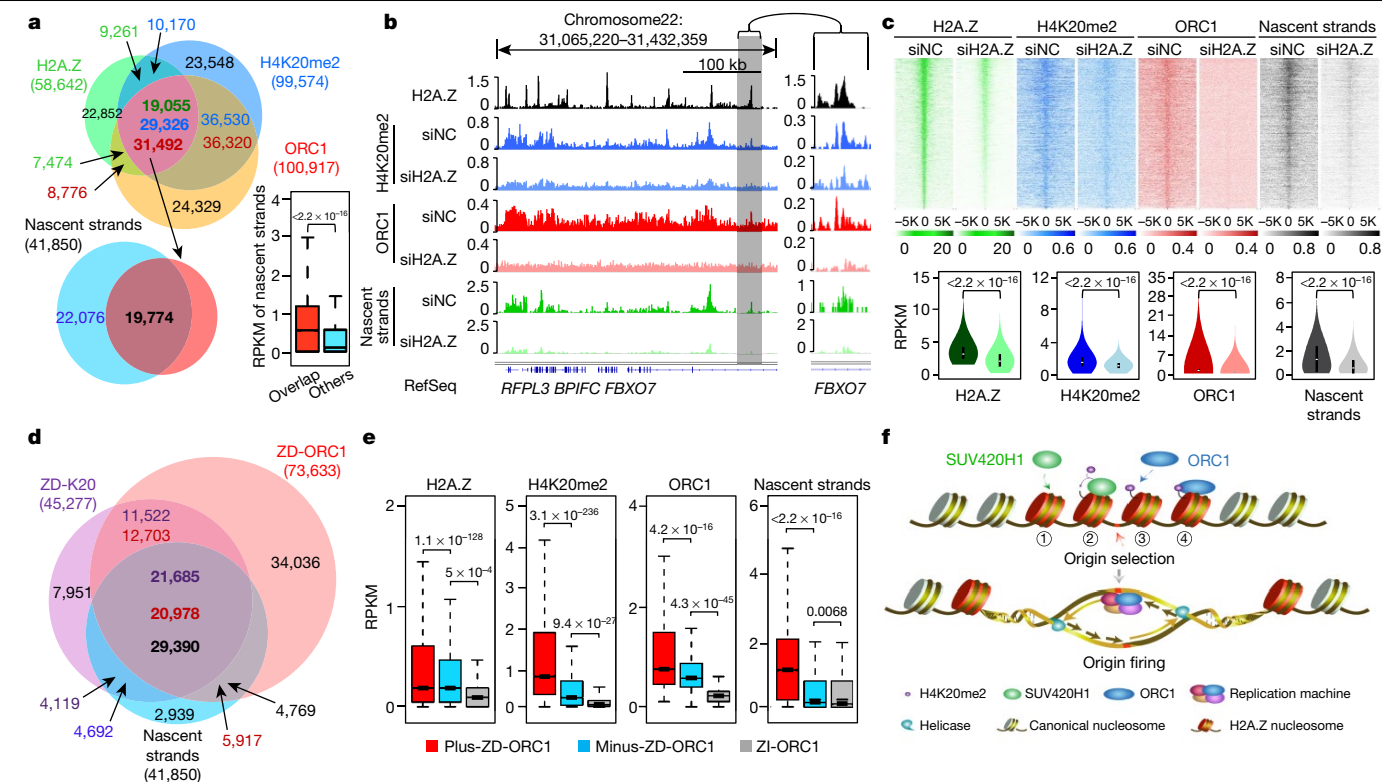
## H2A.Z controls replication origin firing

To investigate how H2A.Z regulates H4K20me2 deposition and ORC1 binding at the genome-wide level, we used chromatin immunoprecipitation with DNA sequencing (ChIP-seq) to map 58,642, 99,574 and 100,917 peaks for H2A.Z, H4K20me2 and ORC1, respectively, in HeLa cells (Fig. 3a). The H4K20me2 ChIP-seq signal was markedly reduced after both SUV420H1 and SUV420H2 were knocked down (Extended Data Fig. 5a). H4K20me1 and H4K20me3 partially overlap with H4K20me2 (Extended Data Fig. 5b), and their ChIP-seq signal increased slightly at the H4K20me2 peak regions after H2AFZ knockdown (Extended Data Fig. 5c). These results validate the specificity of our H4K20me2 ChIP-seq data. The peaks of H2A.Z, H4K20me2 and ORC1 overlap highly with each other (Fig. 3a), and both H4K20me2 and ORC1 levels correlate positively with the H2A.Z level (Extended Data Fig. 5d, e). Moreover, after H2AFZ knockdown, the H4K20me2 and ORC1 levels decrease noticeably at the H2A.Z peak regions that overlap with both H4K20me2 and ORC1 (Fig. 3b, c). Mass-spectrometry analysis of H4K20me2 after H2AFZ

knockdown validates the global decrease in H4K20me2 (Extended Data Fig. 5f). These results support the critical role of H2A.Z in regulating H4K20me2 and ORC1 levels genome wide. Given that ORC1 is degraded during S phase<sup>17</sup>, we analysed ORC1 binding to chromatin in G1/S-phase-arrested cells. The results show that the chromatin fraction of ORC1—and levels of SUV420H1, MCM2 and H4K20me2—decreased after H2AFZ knockdown (Extended Data Fig. 5g), excluding an effect of cell-cycle change on ORC1 binding after H2AFZ knockdown.

To investigate whether H2A.Z regulates the firing of replication origins, we mapped the active replication origins in HeLa cells by nascent-strand sequencing (NS-seq)<sup>18</sup>. We found that treatment with RNase A markedly reduced the nascent-strand signal (Extended Data Fig. 5h). Genome-wide, we detected 41,850 nascent-strand peaks (normalized by the nascent-strand signal following treatment with RNase A), 47.2% of which co-localized with H2A.Z, H4K20me2 and ORC1 simultaneously (Fig. 3a). Moreover, this group of nascent-strand peaks had a higher read density than other nascent-strand peaks (Fig. 3a). The nascent-strand signal also decreased notably after H2AFZ knockdown (Fig. 3b, c), suggesting that H2A.Z is essential for origin firing.

To further analyse the regulatory role of H2A.Z in origin firing, we defined the H2A.Z-regulated H4K20me2 peaks as 'ZD-K20' (H2A.Z-dependent-H4K20me2) (Extended Data Fig. 5i), and the remaining H4K20me2 peaks as 'ZI-K20' (H2A.Z-independent-H4K20me2). We also defined 'ZD-ORC1' (H2A.Z-dependent-ORC1) and 'ZI-ORC1' (H2A.Z-independent-ORC1) (Extended Data Fig. 5j). Remarkably, we found that 73.3% of ZD-K20 peaks overlap with 45.7% of ZD-ORC1 peaks, and that 70.2% of the nascent-strand peaks overlap with ZD-K20 and ZD-ORC1 simultaneously (Fig. 3d). We further defined ZD-ORC1 peaks that



**Fig. 3 | H2A.Z regulates the recruitment of ORC1 to chromatin in order to license origins with higher firing efficiency.** **a**, The top Venn diagram shows the overlap among H2A.Z, H4K20me2 and ORC1 ChIP-seq peaks. The bottom Venn diagram shows the overlap between the 31,492 ORC1 peaks and nascent-strand peaks. The box plot shows the nascent-strand signal in the nascent-strand peaks that overlap the H2A.Z, H4K20me2 and ORC1 peaks ('overlap',  $n = 19,774$ ) and the other nascent-strand peaks ('others',  $n = 22,076$ ). RPKM, reads per kilobase per million mapped reads. **b**, Genome tracks show the signals of H4K20me2, ORC1 and nascent strands. *RFPL3*, *BPIFC* and *FBX07* are RefSeq gene names. **c**, Heat maps and violin plots show the signals of H2A.Z, H4K20me2, ORC1 and nascent strands around the centres of H2A.Z peaks that overlap both H4K20me2 and ORC1 ( $n = 19,055$ ). **d**, Venn diagram showing the overlap among ZD-K20, ZD-ORC1 and nascent-strand regions. **e**, Box plots showing the signals of H2A.Z, H4K20me2, ORC1 and nascent strands in plus-ZD-ORC1, minus-ZD-ORC1 and ZI-ORC1 regions.

overlap ZD-K20 and nascent-strand peaks as 'plus-ZD-ORC1' (the 20,978 ZD-ORC1 peaks in Fig. 3d), which represents those replication origins that are regulated by the H2A.Z-SUV420H1-H4K20me2-ORC1 axis, and the remaining ZD-ORC1 peaks as 'minus-ZD-ORC1'. We found that H2A.Z, H4K20me2, ORC1 and nascent-strand signals were all higher around plus-ZD-ORC1 peaks than around minus-ZD-ORC1 or ZI-ORC1 peaks (Fig. 3e). These results support the essential role of H2A.Z in activating replication origins with a higher firing efficiency through the SUV420H1-H4K20me2-ORC1 axis.

Analysis of the genome-wide distribution of plus-ZD-ORC1 showed that some of these replication origins were located at promoter regions (Extended Data Fig. 5k). In addition, we analysed the DNA sequence features in the 20,978 plus-ZD-ORC1 peaks. Consistent with previous observations<sup>19,20</sup>, we found that G/C-rich and asymmetric A/T-rich motifs were significantly ( $P < 0.01$ ) enriched (Extended Data Table 1), indicating cooperative regulation of origin selection and firing by genetic and epigenetic elements.

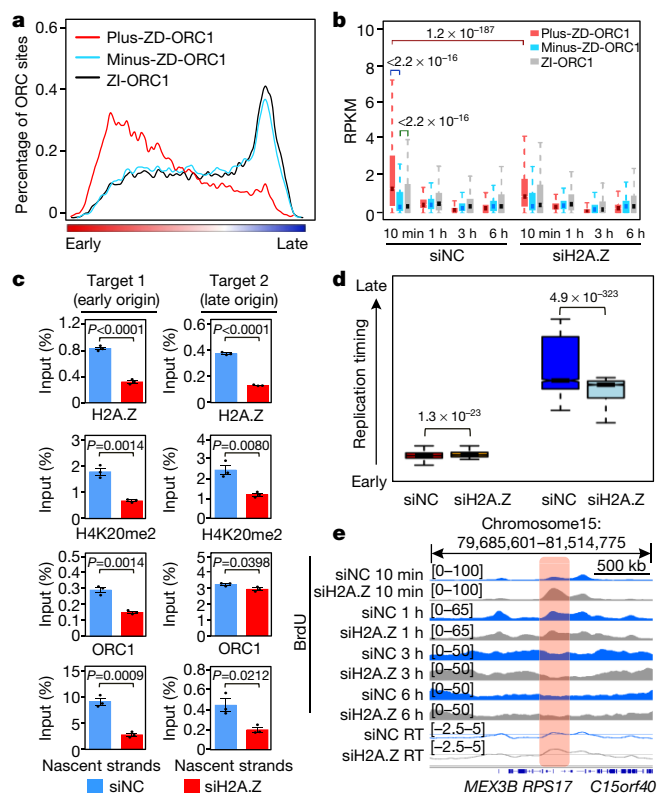
## H2A.Z regulates early-replication origins

To investigate whether the origins regulated by H2A.Z show any preferred timing for replication, we performed BrdU-seq to map

ZD-ORC1 ( $n = 20,978$ ), minus-ZD-ORC1 ( $n = 52,655$ ) and ZI-ORC1 ( $n = 27,284$ ) regions. **f**, Working model. Top, origin selection: H2A.Z nucleosomes bind SUV420H1 directly (step 1) to establish H4K20me2 on chromatin (step 2), which then recruits ORC1 (step 3) to bind to replication origins (red mark; step 4). Bottom, origin firing: the H2A.Z-SUV420H1-H4K20me2-ORC1 axis selectively licenses and activates early replication origins. The experiment in **b** was independently repeated two times with similar results. Data in **a**, **c**, **e** were analysed by two-tailed Wilcoxon test;  $P$ -values are shown in the box plots. For the box plots in **a** and **e**, the centre line represents the median, the box limits are the 25th and 75th percentiles, and the whiskers are the minimum to maximum values. For the violin plots in **c**, the box centres represent the median, the box limits are the 25th and 75th percentiles, and the upper and lower limits make the 95% confidence interval.

replication timing in HeLa cells. As shown in Extended Data Fig. 6a, about 80% of the early-replication domains that we identified overlap with those defined previously by BrdU Repli-seq<sup>21</sup>. We also found that plus-ZD-ORC1 peaks were preferentially enriched at early-replication domains, whereas both minus-ZD-ORC1 and ZI-ORC1 peaks were more enriched at late-replication domains (Fig. 4a). Moreover, plus-ZD-ORC1 peaks were more enriched at the centres of early-replication domains than were the other ORC1 peaks (Extended Data Fig. 6b). Of note, the 10-min BrdU signal (for cells labelled for 10 min immediately after being released from G1/S arrest) in plus-ZD-ORC1 is higher than in other ORC1 peaks, and it decreased obviously after *H2AFZ* knockdown (Fig. 4b). These results suggest that the H2A.Z-SUV420H1-H4K20me2-ORC1 axis preferentially licenses and activates early-replication origins—a conclusion supported by real-time PCR analyses of H4K20me2, ORC1 and nascent strands (Fig. 4c and Extended Data Fig. 6c). Next we analysed dynamic changes in replication timing after *H2AFZ* knockdown, finding that, although the replication timing of early-replication domains did not change substantially, the timing of late-replication regions was advanced (Fig. 4d, e and Extended Data Fig. 6d, e). We then arrested H2A.Z-depleted cells at the G1/S boundary. After release, H2A.Z-depleted cells progressed through S phase without any defects (Extended Data Fig. 7a, b), suggesting that, after *H2AFZ* knockdown,





**Fig. 4 | H2A.Z regulates early replication origins and replication timing.** **a**, Graph showing the replication timing of plus-ZD-ORC1 ( $n = 20,978$ ), minus-ZD-ORC1 ( $n = 52,655$ ) and ZI-ORC1 ( $n = 27,284$ ) peaks. **b**, Box plot showing the BrdU signal in siNC or siH2A.Z cells. **c**, Real-time PCR analysis of the ChIP signal from H2A.Z, H4K20me2 and ORC1, and nascent-strand signals in siNC or siH2A.Z cells. **d**, Box plot showing the dynamics of replication timing of the 2,000 earliest and 300 latest replication origins (nascent-strand peaks). **e**, Genome tracks of a relatively late replication domain (red shaded area) show the increased BrdU signal at 10 min and advanced replication timing after *H2AFZ* knockdown. Numbers in square brackets indicate the data range of corresponding track. *MEX3B*, *C15orf40* and *RPS17* are RefSeq gene names. Data in **c** are mean  $\pm$  s.e.m.;  $n = 3$  biological replicates; two-tailed unpaired *t*-test. Data in **b**, **d** were analysed by two-tailed Wilcoxon test. *P*-values are indicated within **b**–**d**. The experiment in **e** was independently repeated twice with similar results.

replication timing was reprogrammed to ensure that the whole genome was replicated efficiently.

We next investigated whether the R257 and K333 residues of SUV420H1 are involved in activating early origins in vivo. We found that neither SUV420H1<sub>R257A</sub> nor SUV420H1<sub>K333A</sub> could rescue H4K20me2 after knocking down endogenous *SUV420H1* (Extended Data Fig. 7c). In addition, through real-time PCR we found that these two SUV420H1 mutants could not rescue nascent-strand or BrdU signals (labelled for 10 minutes immediately after release from G1/S arrest) (Extended Data Fig. 7d). Thus, H2A.Z regulates the selection and activation of early-replication origins through SUV420H1.

## H2A.Z in activated T cells

To study the function of H2A.Z-regulated replication in a more physiological context, we conditionally knocked out (CKO) *H2az1/H2az2* in T cells by generating CD4<sup>Cre</sup>H2A.Z<sup>f/f</sup> mice (Extended Data Fig. 8a). We found that the number of mature T cells in the spleen decreased dramatically in H2A.Z CKO mice (Extended Data Fig. 8b, c). H2A.Z CKO T cells showed reduced BrdU incorporation in the homeostatic state, and reduced dilution of carboxyfluorescein succinimidylester

(CFSE) in activated T cells on stimulation with CD3 and CD28 antibodies (Extended Data Fig. 8d–f). In addition, fluorescence-activated cell sorting (FACS) analysis showed that activated T cells from CKO mice had a prolonged G1 phase and a shorter S phase than the wild type (Extended Data Fig. 8g, h). These results suggest an essential role of H2A.Z in DNA replication and T-cell proliferation.

ChIP-seq analysis showed that the H4K20me2 and nascent-strand signals were markedly reduced in H2A.Z CKO activated T cells (Extended Data Fig. 8i). Western blot analysis confirmed the global loss of H4K20me2 upon H2A.Z depletion in activated T cells (Extended Data Fig. 8j). We classified nascent-strand signals into ZD-NS (H2A.Z-dependent-nascent strand;  $n = 18,382$ ) and ZI-NS (H2A.Z-independent-nascent strand;  $n = 7,901$ ). We found that H4K20me2, as well as H2A.Z and nascent strands, in ZD-NS regions were higher than those in ZI-NS regions (Extended Data Fig. 8k), indicating that, as with HeLa cells, the H2A.Z–SUV420H1–H4K20me2–ORC1 axis has an important regulatory role in the selection and firing of replication origins in activated T cells.

## Discussion

We have shown that nucleosomes comprising H2A.Z histones can directly bind SUV420H1 to efficiently stimulate the dimethylation of H4 K20 residues, thereby licensing and activating early-replication origins (Fig. 3f). However, given that H2A.Z nucleosomes are much less abundant than canonical H2A nucleosomes in cells, it remains unknown how H2A.Z regulates the global abundance of H4K20me2 in vivo. It has been shown that H4K20me2 levels are passively diluted twofold during DNA replication, recovering gradually by the next G1 phase<sup>22</sup>. Notably, H2A.Z is lost on nascent chromatin after DNA replication, and is restored along with H4K20me2 (ref. <sup>22</sup>), suggesting that it has an essential role in establishing H4K20me2 on newly synthesized histones. Recently, a class of *cis*-regulatory elements called early-replicating control elements was found to be essential for maintaining the timing of early replication<sup>23</sup>. Another study showed that poly(dA:dT) tracts were associated with efficient replication origins<sup>20</sup>. Interestingly, nucleosomes were depleted at the centres of these poly(dA:dT) tracts, but strongly positioned at flanking regions<sup>20</sup>, akin to the features of chromatin structure seen at replication origins in yeast<sup>24</sup>. We have found here that active early origins are highly enriched with G/C-rich and asymmetric A/T-rich motifs. Thus, we speculate that coordination between genetic determinants and epigenetic features is involved in fine-tuning the licensing and activation of replication origins during the cell cycle.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-019-1877-9>.

1. Fragkos, M., Ganier, O., Coulombe, P. & Méchali, M. DNA replication origin activation in space and time. *Nat. Rev. Mol. Cell Biol.* **16**, 360–374 (2015).
2. MacAlpine, D. M. & Almouzni, G. Chromatin and DNA replication. *Cold Spring Harb. Perspect. Biol.* **5**, a010207 (2013).
3. O'Donnell, M., Langston, L. & Stillman, B. Principles and concepts of DNA replication in bacteria, archaea, and eukarya. *Cold Spring Harb. Perspect. Biol.* **5**, a010108 (2013).
4. Dhar, M. K., Sehgal, S. & Kaul, S. Structure, replication efficiency and fragility of yeast ARS elements. *Res. Microbiol.* **163**, 243–253 (2012).
5. Kuo, A. J. et al. The BAH domain of ORC1 links H4K20me2 to DNA replication licensing and Meier-Gorlin syndrome. *Nature* **484**, 115–119 (2012).
6. Pesavento, J. J., Bullock, C. R., LeDuc, R. D., Mizzen, C. A. & Kelleher, N. L. Combinatorial modification of human histone H4 quantitated by two-dimensional liquid chromatography coupled with top down mass spectrometry. *J. Biol. Chem.* **283**, 14927–14937 (2008).
7. Cayrou, C. et al. The chromatin environment shapes DNA replication origin organization and defines origin classes. *Genome Res.* **25**, 1873–1885 (2015).
8. Costas, C. et al. Genome-wide mapping of *Arabidopsis thaliana* origins of DNA replication and their associated epigenetic marks. *Nat. Struct. Mol. Biol.* **18**, 395–400 (2011).

9. Natsume, T., Kiyomitsu, T., Saga, Y. & Kanemaki, M. T. Rapid protein depletion in human cells by auxin-inducible degron tagging with sShort homology donors. *Cell Reports* **15**, 210–218 (2016).
10. Ohta, S., Tatsumi, Y., Fujita, M., Tsurimoto, T. & Obuse, C. The ORC1 cycle in human cells: II. Dynamic changes in the human ORC complex during the cell cycle. *J. Biol. Chem.* **278**, 41535–41540 (2003).
11. Schotta, G. et al. A chromatin-wide transition to H4K20 monomethylation impairs genome integrity and programmed DNA rearrangements in the mouse. *Genes Dev.* **22**, 2048–2061 (2008).
12. Jørgensen, S., Schotta, G. & Sørensen, C. S. Histone H4 lysine 20 methylation: key player in epigenetic regulation of genomic integrity. *Nucleic Acids Res.* **41**, 2797–2806 (2013).
13. Fan, J. Y., Rangasamy, D., Luger, K. & Tremethick, D. J. H2A.Z alters the nucleosome surface to promote HP1 $\alpha$ -mediated chromatin fiber folding. *Mol. Cell* **16**, 655–661 (2004).
14. Wang, Y. et al. Histone variants H2A.Z and H3.3 coordinately regulate PRC2-dependent H3K27me3 deposition and gene expression regulation in mES cells. *BMC Biol.* **16**, 107 (2018).
15. Wu, H. et al. Crystal structures of the human histone H4K20 methyltransferases SUV420H1 and SUV420H2. *FEBS Lett.* **587**, 3859–3868 (2013).
16. Suto, R. K., Clarkson, M. J., Tremethick, D. J. & Luger, K. Crystal structure of a nucleosome core particle containing the variant histone H2A.Z. *Nat. Struct. Biol.* **7**, 1121–1124 (2000).
17. Tatsumi, Y., Ohta, S., Kimura, H., Tsurimoto, T. & Obuse, C. The ORC1 cycle in human cells: I. Cell cycle-regulated oscillation of human ORC1. *J. Biol. Chem.* **278**, 41528–41534 (2003).
18. Cayrou, C., Grégoire, D., Coulombe, P., Danis, E. & Méchali, M. Genome-scale identification of active DNA replication origins. *Methods* **57**, 158–164 (2012).
19. Bartholdy, B., Mukhopadhyay, R., Lajugie, J., Aladjem, M. I. & Bouhassira, E. E. Allele-specific analysis of DNA replication origins in mammalian cells. *Nat. Commun.* **6**, 7051 (2015).
20. Tubbs, A. et al. Dual roles of poly(dA:dT) tracts in replication initiation and fork collapse. *Cell* **174**, 1127–1142 (2018).
21. Dellino, G. I. et al. Genome-wide mapping of human DNA-replication origins: levels of transcription at ORC1 sites regulate origin selection and replication timing. *Genome Res.* **23**, 1–11 (2013).
22. Alabert, C. et al. Two distinct modes for propagation of histone PTMs across the cell cycle. *Genes Dev.* **29**, 585–590 (2015).
23. Sima, J. et al. Identifying cis elements for spatiotemporal control of mammalian DNA replication. *Cell* **176**, 816–830 (2019).
24. Eaton, M. L., Galani, K., Kang, S., Bell, S. P. & MacAlpine, D. M. Conserved nucleosome positioning defines replication origins. *Genes Dev.* **24**, 748–753 (2010).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2019



## Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

## Data availability

ChIP-seq, NS-seq, BrdU-seq and RNA-seq data have been deposited in the Gene Expression Omnibus (GEO) under accession number GSE134988. Further information and resources and reagents are available from G.L. upon reasonable request. Source data for Figs. 1, 2, 4, Extended Data Figs. 1–3, 5–8 and Supplementary Figures are provided with the paper.

## Code availability

Custom codes used for data analysis in this paper can be found at <https://github.com/Leavy-Zhang/RepliCode>.

**Acknowledgements** We thank M. Méchali for advice on nascent-strand sequencing; and B. Stillman and Z. Zhang for critical reading and discussion of the manuscript. This work was supported by grants to G.L. from the Ministry of Science and Technology of China (2017YFA0504202), the National Natural Science Foundation of China (31525013, 31630041 and 31521002), and the Chinese Academy of Sciences (CAS) Strategic Priority Research Program (XDB19040202). The work was also supported by the CAS Key Research Program on Frontier Science (QYZDY-SSW-SMC020) and a Howard Hughes Medical Institute (HHMI) international research scholar grant (55008737) to G.L. This work was also supported by a

grant from the National Natural Science Foundation of China to L.Z. (31801062); the China Postdoctoral Science Foundation to Z.W. (2019M650871); the Ministry of Science and Technology of China (2018YFE0203302) and the National Natural Science Foundation of China (31871290) to P.C.; and the Ministry of Science and Technology of China to J.P. (2016YFA05023032). All fluorescence imaging data were collected at the Center for Bio-imaging, Core Facility for Protein Sciences (Institute of Biophysics, CAS). All radioactivity experiments were performed at the radioactive isotope laboratory (Institute of Biophysics, CAS), with guidance from H. J. Zhang in handling radioactive materials.

**Author contributions** H.L. carried out experiments and composed the figures. L.Z. performed bioinformatics analysis of next-generation sequencing data and generated figures. Z.W. analysed the phenotypic effects of the H2A.Z-knockdown on cell-cycle deficiency, and assisted with ChIP-seq experiments and the analysis of next-generation sequencing data. M.L. analysed the phenotype of H2A.Z CKO mice and provided T cells for biochemistry experiments. W.Z. and H.D. performed mass-spectrometry analysis of the H4K20 modification of histone methyltransferase products, mononucleosomes in vivo and cells treated with siNC or siHA.Z oligonucleotides. X.C. and F.Y. performed mass-spectrometry analysis of H2A.Z nucleosome binding partners. P.Z. and X.W. assisted with the quantification of the histone methyltransferase reaction, plasmid construction, and protein purification. T.L. and J.P. performed the in silico analysis of the interaction between H2A.Z and SUV420H1. L.C. assisted with the analysis of H2A.Z and H4K20me2 sequencing data. C.J. constructed the H2A.Z CKO mice and assisted with their phenotype analysis. G.W. assisted with plasmid construction and protein purification. P.C. and R.M. helped to discuss the project. M.Z. conceived the project on the phenotype of H2A.Z CKO mice. G.L. conceived and supervised the project, analysed the data and wrote the manuscript.

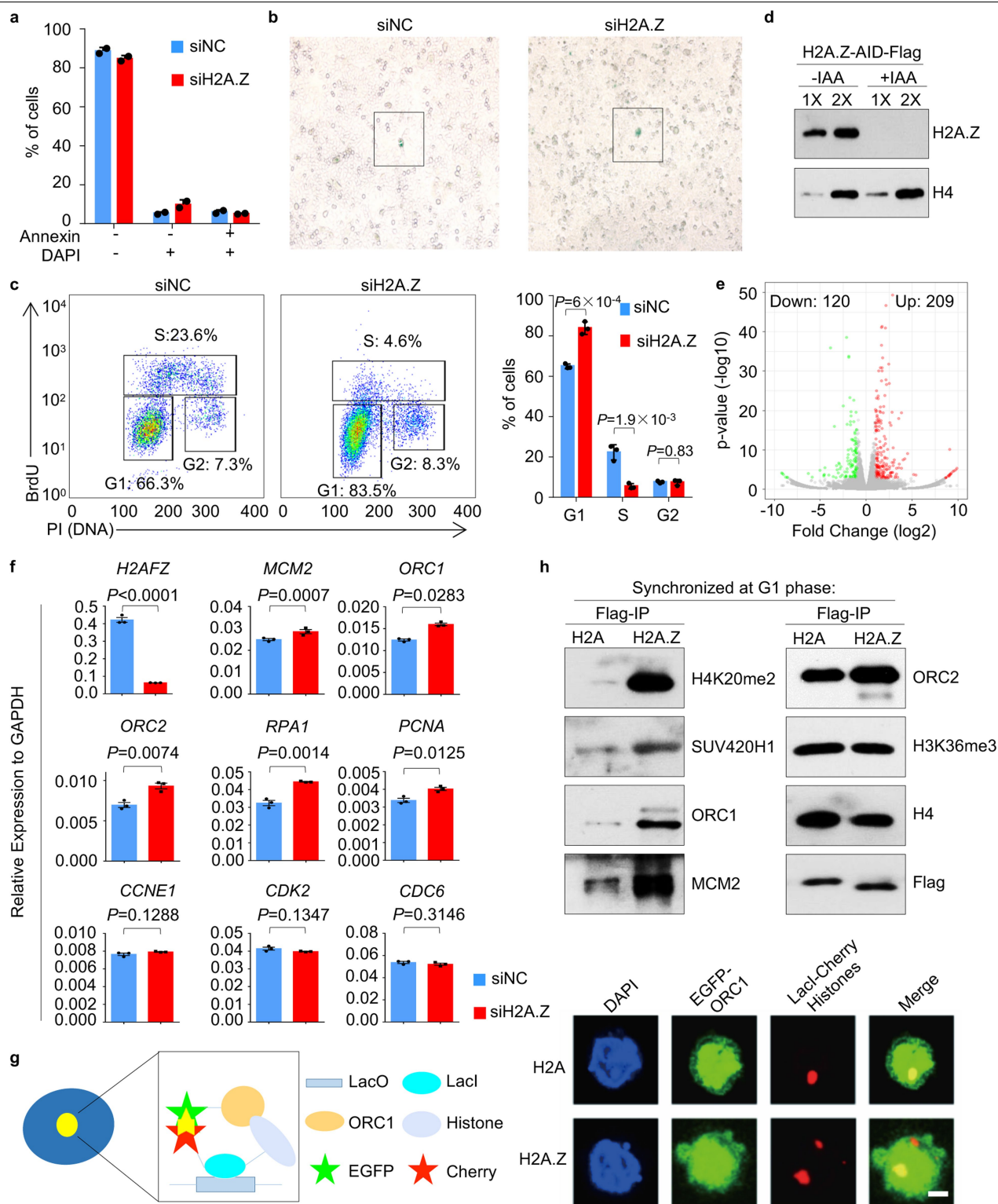
**Competing interests** The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41586-019-1877-9>.

**Correspondence and requests for materials** should be addressed to M.Z. or G.L.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>.

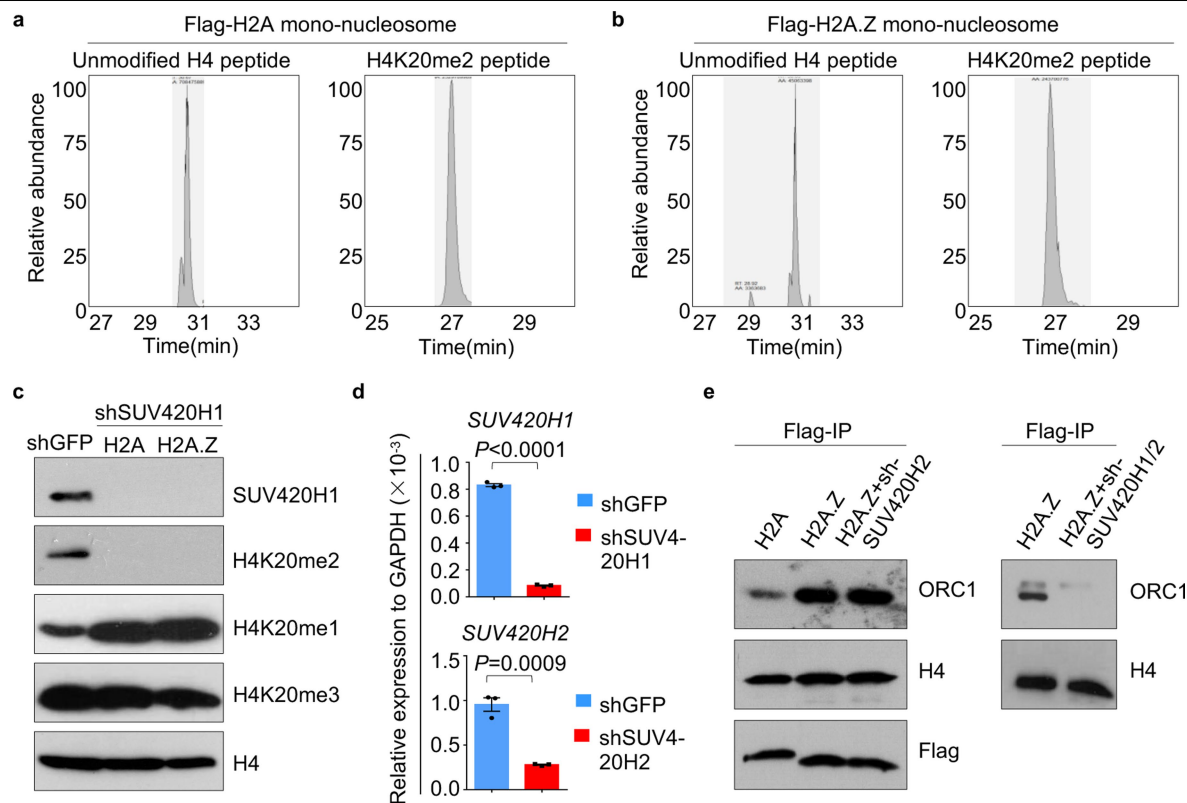


**Extended Data Fig. 1** | See next page for caption.

## Extended Data Fig. 1 | Analysis of phenotypes and gene expression after *H2AFZ* knockdown, and analysis of the interaction between H2A.Z and ORC1.

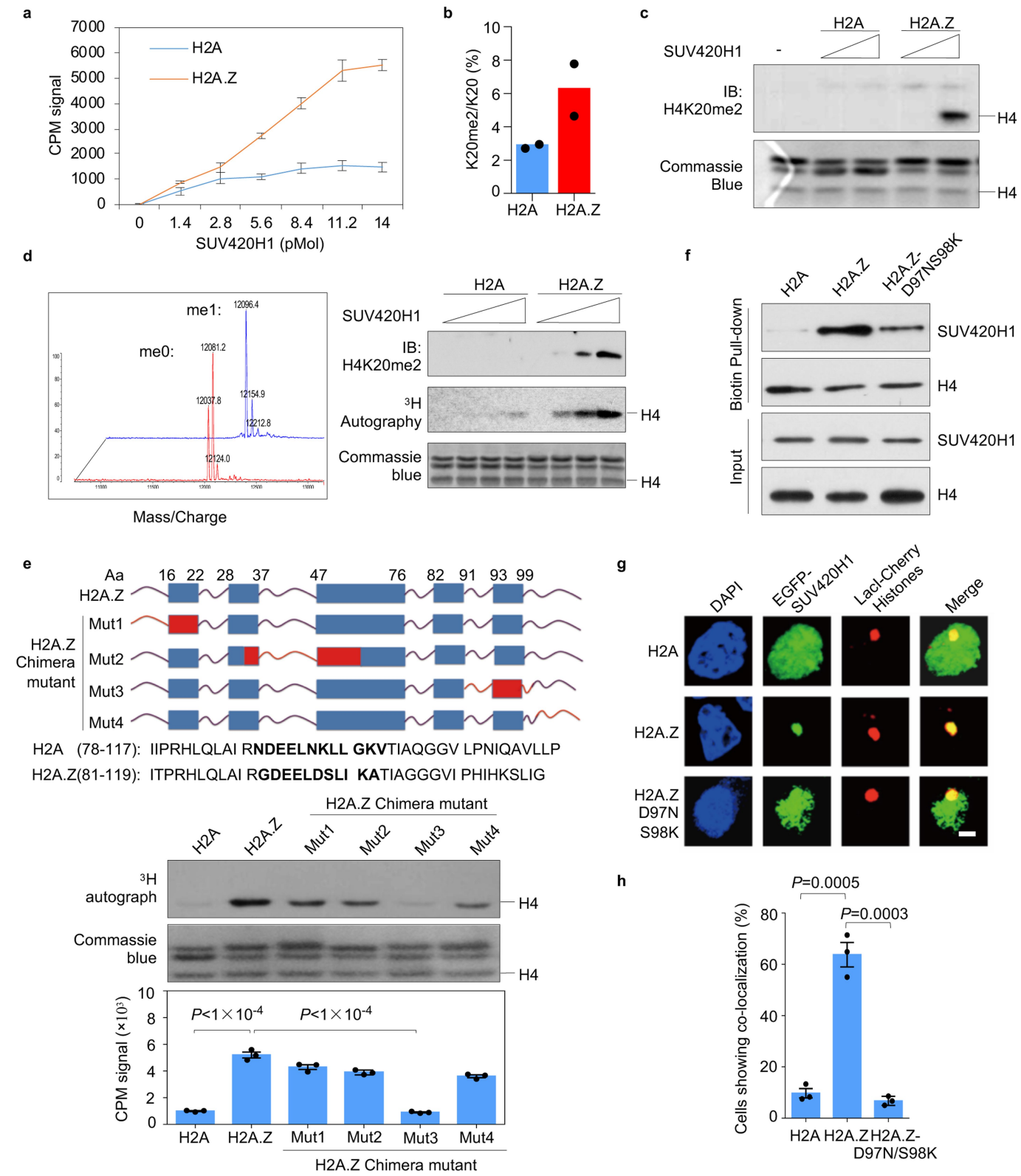
**a**, Statistical outcome of FACS analysis of apoptosis of cells treated with siNC or siH2A.Z oligonucleotides. Annexin is used as a marker of apoptotic cells; DAPI is 4',6-diamidino-2-phenylindole, a nuclear marker. **b**, Representative images of cells undergoing senescence (within boxed regions), shown by  $\beta$ -galactosidase staining. **c**, Cell-cycle analysis of siNC or siH2A.Z HeLa cells. Cells were pulse labelled with BrdU and stained with propidium iodide (PI), and then analysed by FACS. **d**, Western blots showing the H2A.Z level when cells are released from arrest at G2/M phase after treatment with IAA. AID, auxin-inducible degron. **e**, Volcano plot showing genome-wide expression dynamics after *H2AFZ* knockdown;  $n = 33,835$ . Genes with a  $\log_2$ (fold change) of 1 or more and  $P$ -values of less than 0.01 were selected as differentially expressed genes. **f**, Real-time PCR analysis of gene expression in siNC or siH2A.Z cells. The expression level was normalized to that of the glyceraldehyde-3-phosphate dehydrogenase gene (*GAPDH*). **g**, Left, diagram showing the LacO/LacI

targeting system in A03\_1 cells. An interaction between a LacI-Cherry-tagged histone and an enhanced green fluorescent protein (EGFP)-tagged ORC protein would result in an overlap of green fluorescence with red fluorescence. Right, there is no interaction in this experiment between ORC1 and H2A or H2A.Z. Scale bar, 5  $\mu\text{m}$ . **h**, Western blot analysis of the enrichment of H4K20me2, H3K36me3 (negative control), SUV420H1, ORC1, ORC2 and MCM2 on H2A or H2A.Z nucleosomes from G1-phase-synchronized cells. Data in panel **a** are means;  $n = 2$  biological replicates, with dot plots overlaid. Data in panels **c**, **f** are means  $\pm$  s.e.m.;  $n = 3$  biological replicates; two-tailed unpaired  $t$ -test. The  $\beta$ -galactosidase staining in panel **b**, the FACS results in **c**, the western blots in **d**, **h**, and the fluorescence image in **g** were independently repeated three times with similar results. H4 was used as a loading control and sample processing control in **d** and **h**. For gel source data, see Supplementary Fig. 1. For imaging source data, see Supplementary Fig. 2. For the FACS gating strategy, see Supplementary Fig. 3.



**Extended Data Fig. 2 | H2A.Z interacts with ORC1 in an H4K20me2-dependent manner.** **a, b**, High-performance liquid chromatography (HPLC) of unmodified H4 peptide and H4K20me2 peptide in H2A (**a**) and H2A.Z (**b**) nucleosomes. **c**, Western blot analysis showing the signal of SUV420H1 and H4K20me1/2/3 in shGFP cells or shSUV420H1 cells with stably expressed Flag-tagged H2A or H2A.Z. **d**, Real-time PCR analysis shows the level of expression of *SUV420H1* and *SUV420H2* in cells transfected with shRNA targeting GFP, *SUV420H1* or *SUV420H2*. The expression level was normalized to that of GAPDH. **e**, Left, enrichment of ORC1 on H2A or H2A.Z mononucleosome from

wild-type or *SUV420H2*-knockdown cells. Right, enrichment of ORC1 on H2A.Z mononucleosomes from wild-type or *SUV420H1/2*-knockdown cells. Data in panel **d** are means  $\pm$  s.e.m.;  $n = 3$  biological replicates; two-tailed unpaired *t*-test. The experiment in panel **a** was independently repeated four times with similar results. The experiments in panels **c, e** were independently repeated twice with similar results. H4 was used as a loading control and sample processing control in **c** and **e**, respectively. For gel source data, see Supplementary Fig. 1.

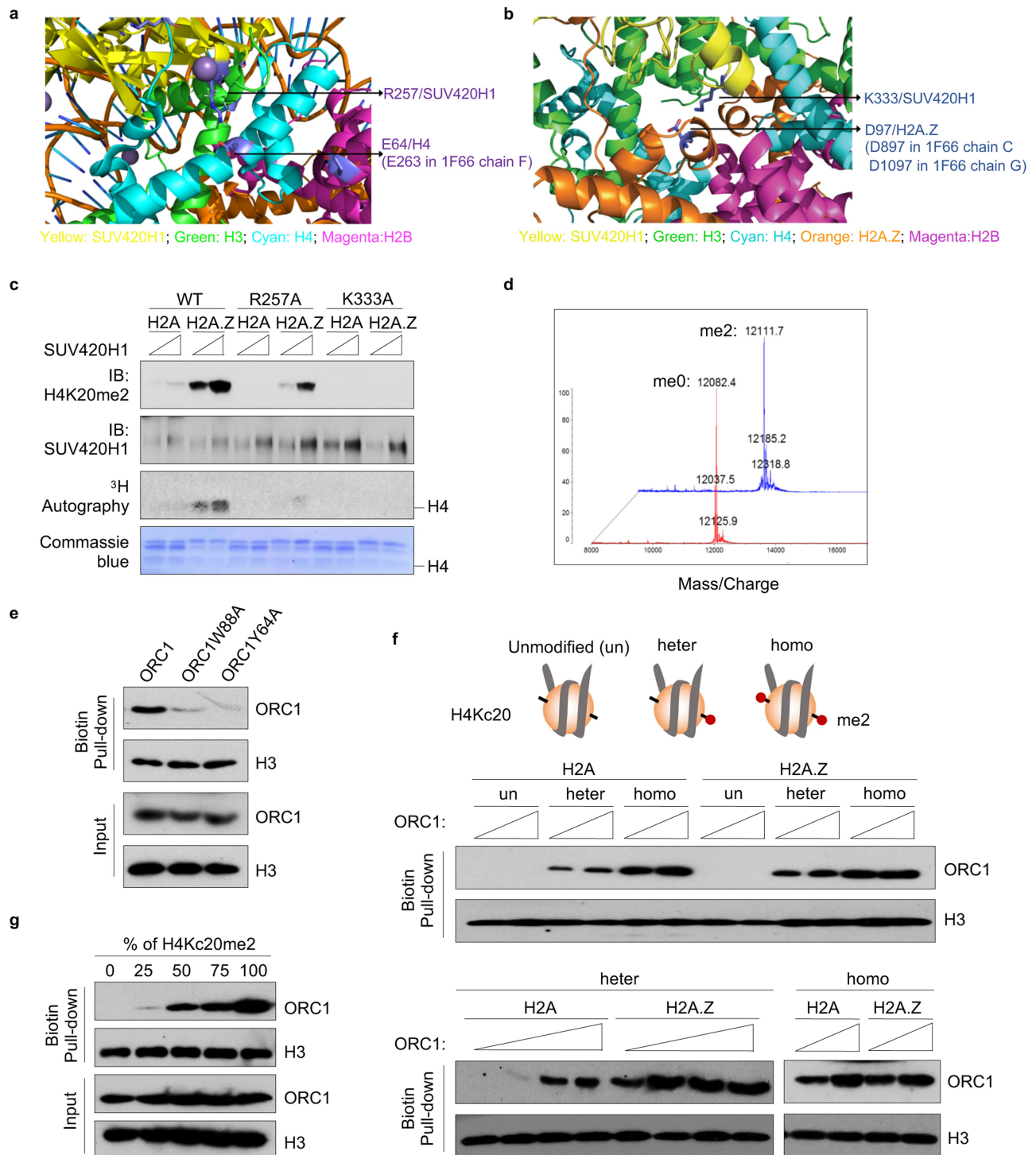


Extended Data Fig. 3 | See next page for caption.



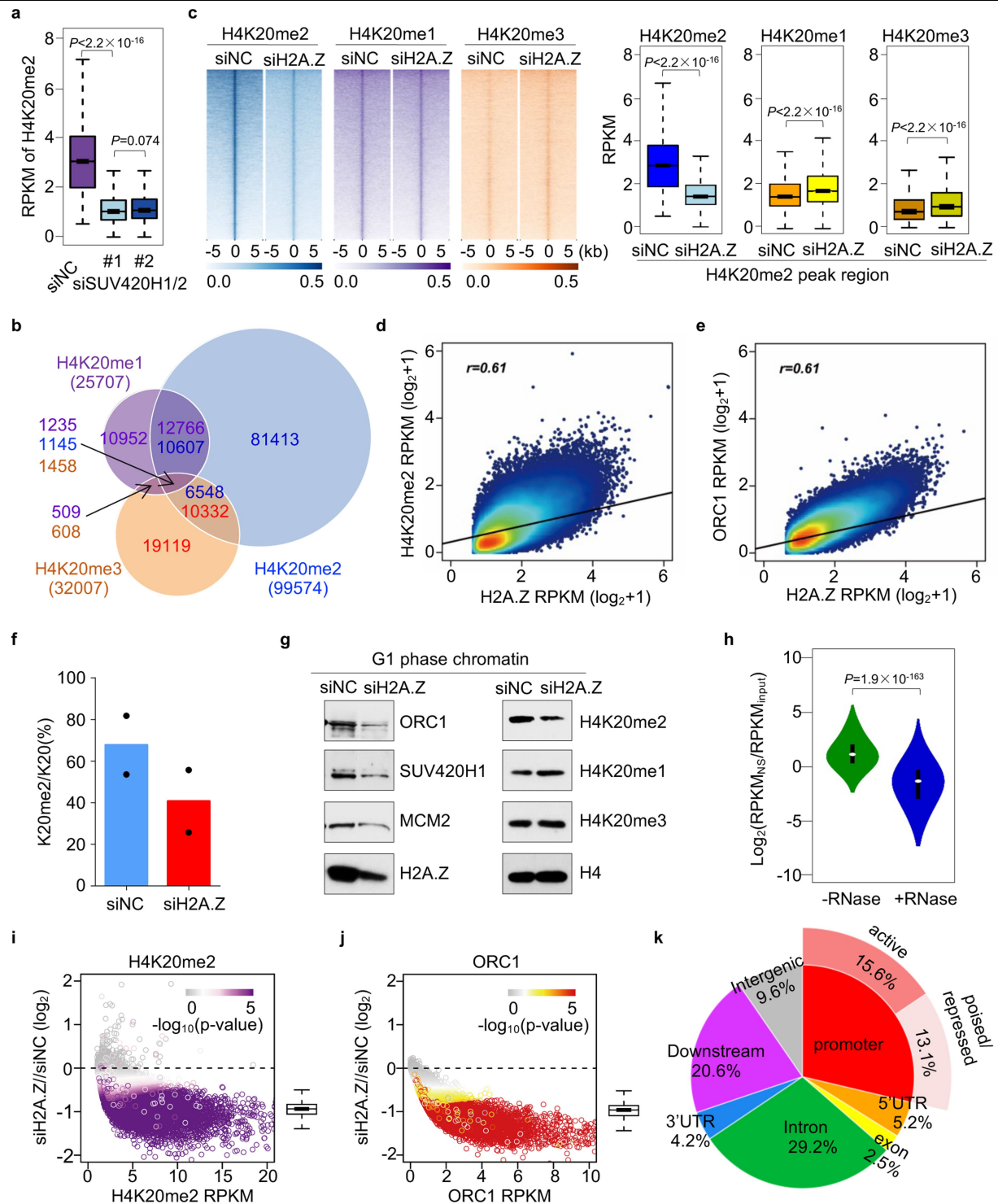
**Extended Data Fig. 3 | H2A.Z enhances the binding of SUV420H1 to promote its enzymatic activity.** **a**, Liquid scintillation results from analysis of SUV420H1 histone methyltransferase activity, using H2A or H2A.Z mononucleosomes as substrates ( $n = 3$  biological replicates). **b**, Mass-spectrometry analysis of H4K20me2 modification by SUV420H1, using H2A or H2A.Z mononucleosomes as substrates. **c**, Western blot analysis of products from histone methyltransferase assay of SUV420H1 using H2A or H2A.Z mononucleosomes as substrates. IB, immunoblot. **d**, Left, mass-spectrometry analysis of monomethylated and unmethylated H4 histones from chemical methylation reactions in vitro. Right, western blot analysis and  $^3\text{H}$  autography show that H2A.Z promotes the activity of SUV420H1 on an H4K<sub>20</sub>me1 substrate. **e**, Upper panel, diagram showing four chimaeric mutants of H2A.Z, with the regions in red replaced with the corresponding regions of H2A. The sequences of the region containing D97 and S98 of H2A.Z and the corresponding region of H2A are shown below the diagram (in single-letter code). Lower panel,  $^3\text{H}$  autograph and liquid scintillation analysis of the methyltransferase activity of SUV420H1 on mononucleosomes containing

H2A.Z chimaeric mutants. **f**, Western blot analysis following the pulldown of biotinylated mononucleosomes shows an interaction between SUV420H1 and mononucleosomes containing wild-type H2A.Z or the H2A.Z<sub>D97N/S98K</sub> mutant. **g**, The interaction between SUV420H1 and H2A, H2A.Z or H2A.Z<sub>D97N/S98K</sub> was analysed by LacO/LacI targeting. Scale bar, 5  $\mu\text{m}$ . **h**, Statistical results from the LacO/LacI targeting assay of panel **g**, showing the percentage of cells in which EGFP-SUV420H1 co-localizes with the indicated histones. Data in panel **a** are means  $\pm$  s.d.;  $n = 3$  biological replicates. Data in panel **b** are means;  $n = 2$  biological replicates, with dot plot overlaid. Data in panels **e**, **h** are means  $\pm$  s.e.m.;  $n = 3$  biological replicates; two-tailed unpaired  $t$ -test. The western blots in panels **c**, **d**, **f**, the  $^3\text{H}$  autography in panels **d**, **e**, the mass spectrometry in panel **d** and the fluorescence imaging in panel **g** were independently repeated three times with similar results. H4 was used as a loading control and sample processing control in panels **c**–**f**. For gel source data, see Supplementary Fig. 1. For imaging source data, see Supplementary Fig. 2.



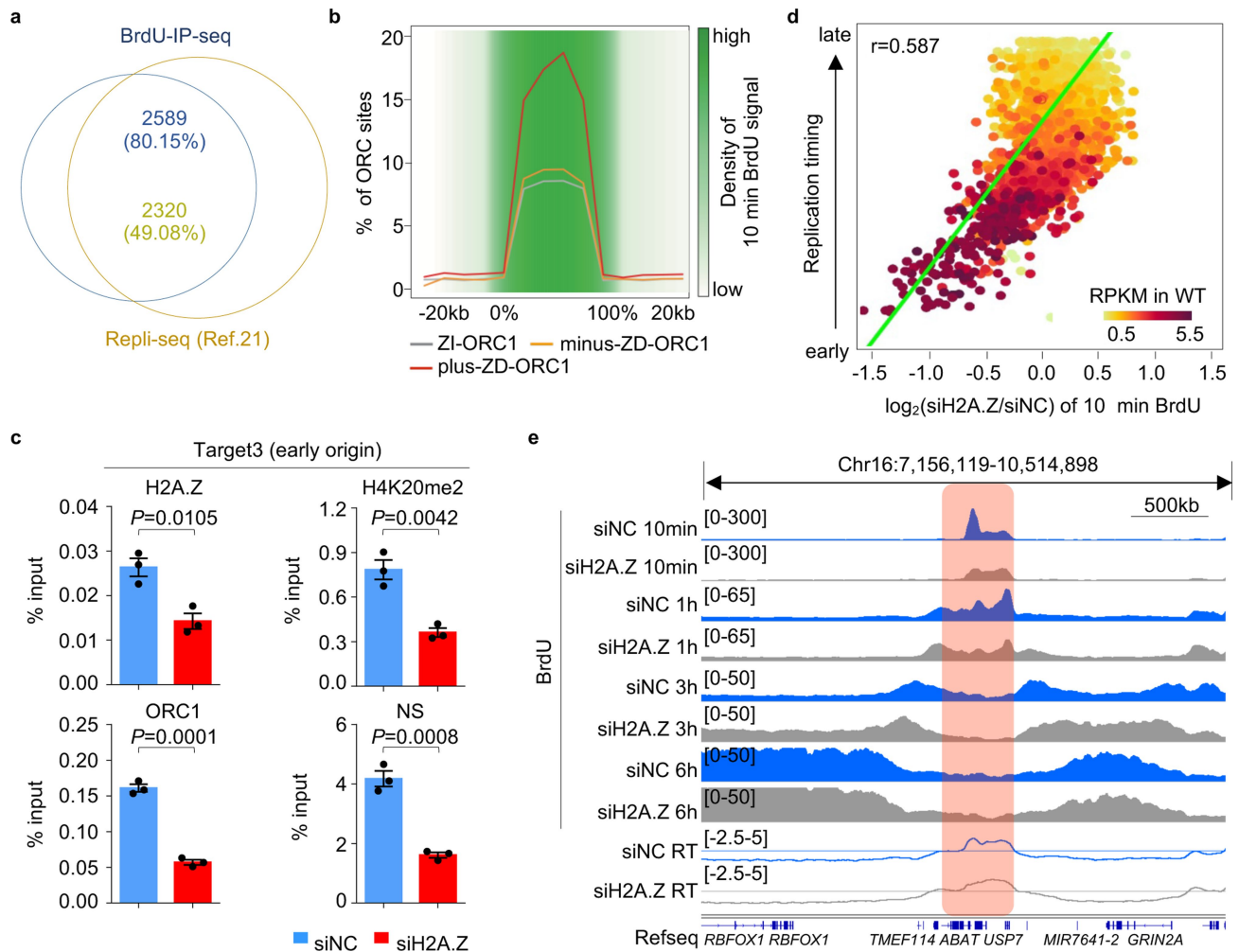
**Extended Data Fig. 4 | H4K20me2 dosage dependent interaction between ORC1 and H4K20me2 nucleosomes. a, b,** Docking of the SUV420H1 and H2A.Z nucleosome structures shows the interaction between R257 of SUV420H1 and E64 of H4 (a), and the interaction between K333 of SUV420H1 and D97 of H2A.Z (b). **c,** Western blots and <sup>3</sup>H autography analysis of the products of a histone methyltransferase assay using wild-type SUV420H1 or SUV420H1 mutants. **d,** Mass-spectrometry analysis of H4 histones with dimethylated K<sub>20</sub>, produced through chemical methylation reactions in vitro; mass spectrometry was performed once to confirm the methylation state. **e,** Western blot analysis

of the interaction between H2A.Z mononucleosomes with the H4K<sub>20</sub>me2 modification and ORC1 or ORC1 BAH-domain mutants. **f,** Western blots show the interaction between ORC1 and H2A or H2A.Z mononucleosomes with different H4K<sub>20</sub>me2 states. **g,** Western blots show the interaction between ORC1 and H2A.Z polynucleosomes with increasing ratios of a H4K<sub>20</sub>me2 octamer. The western blots in panels c, e–g and the <sup>3</sup>H autography in panel c were independently repeated twice with similar results. H4 was used as a loading control in panel c. H3 was used as a loading control in panels e–g. For gel source data, see Supplementary Fig. 1.



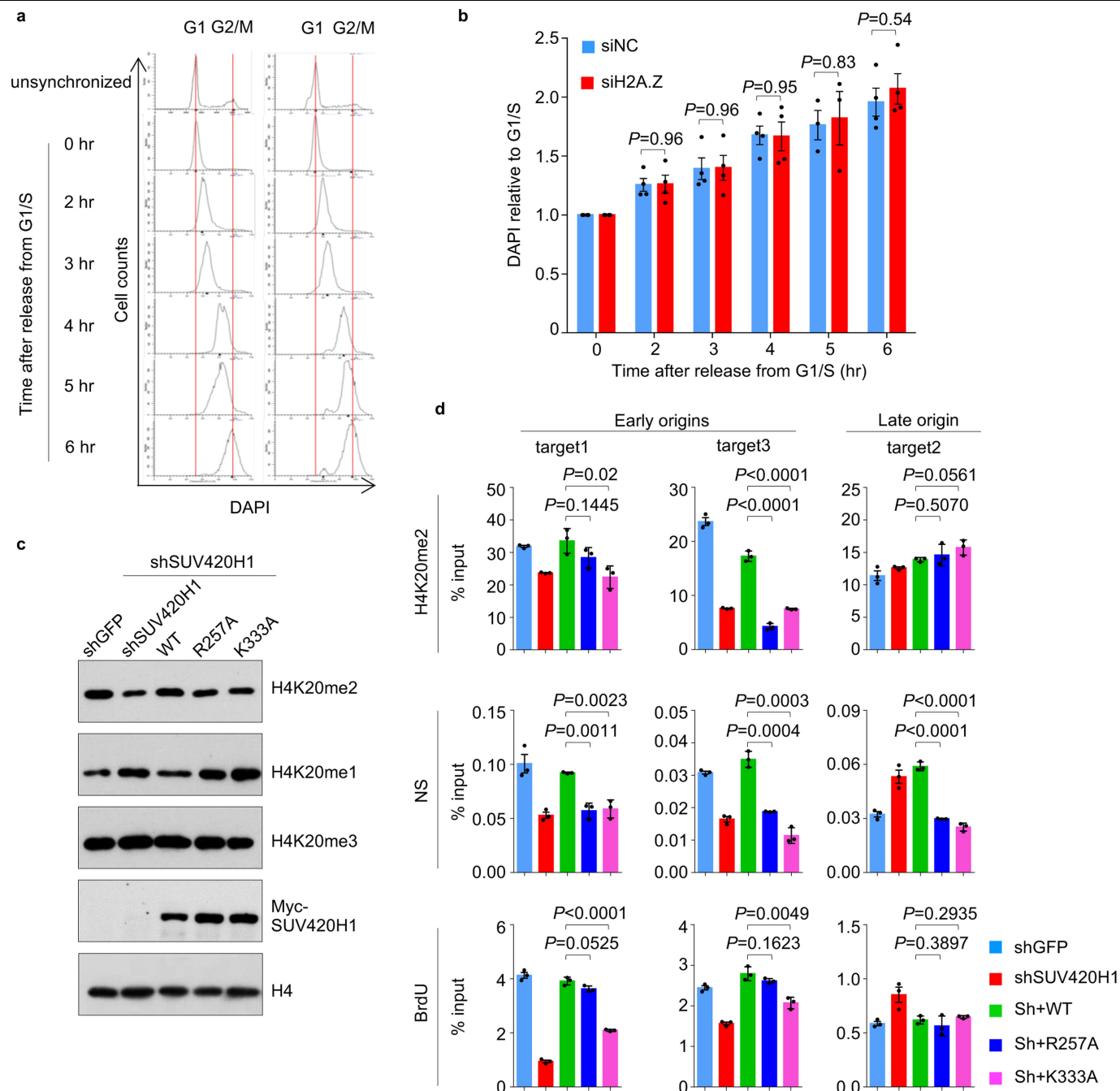
**Extended Data Fig. 5 | H2A.Z regulates H4K20me2 and ORC1 on a genome-wide level.** **a**, Box plot showing the dynamics of H4K20me2 at H4K20me2 peak regions ( $n = 99,574$ ) after knocking down both *SUV420H1* and *SUV420H2*. **b**, Venn diagram showing the overlap between H4K20me1, H4K20me2 and H4K20me3 peaks genome wide. **c**, Heat maps and corresponding box plots showing the dynamics of H4K20me1, H4K20me2 and H4K20me3 at 10-kilobase regions around the centres of the H4K20me2 peaks ( $n = 99,574$ ). **d**, **e**, Dot plot showing a positive correlation between H2A.Z and H4K20me2 (**d**) and H2A.Z and ORC1 (**e**) at H2A.Z peaks ( $n = 58,642$ ).  $r$ , Pearson's correlation coefficient. **f**, Mass-spectrometry analysis of chromatin H4K20me2 abundance after *H2AFZ* knockdown. **g**, Western blot analysis of ORC1, SUV420H1, MCM2, H4K20me1, H4K20me2 and H4K20me3 in the chromatin fraction from cells arrested at G1 phase. **h**, Box plot showing the nascent-strand (NS) signal of RNase-treated and untreated samples from the NS peaks ( $n = 41,850$ ). **i**, **j**, Dot plot showing the

dynamics of H4K20me2 (**i**;  $n = 99,574$ ) and ORC1 (**j**;  $n = 100,917$ ) after knocking down *H2AFZ*. **k**, Genome-wide distribution of plus-ZD-ORC1 peaks ( $n = 20,978$ ). UTR, untranslated region. Data in panels **a**, **c**, **h** were analysed by two-tailed Wilcoxon test. Data in panel **f** are means ( $n = 2$  biological replicates) with dot plots overlaid. Data in panels **i**, **j** were analysed by one-sided Fisher's exact test without adjustments for multiple comparisons. The western blots in panel **g** were independently repeated twice with similar results, and H4 was used as a loading control and sample processing control. For the box plots in panels **a**, **c**, the centre lines represent the medians, the box limits are the 25th and 75th percentiles, and the whiskers are the minimum to maximum values. For the violin plots in panel **h**, the centres of the boxes represent the medians, the box limits are the 25th and 75th percentiles, and the upper and lower limits show the 95% confidence interval. For gel source data, see Supplementary Fig. 1.



**Extended Data Fig. 6 | H2A.Z regulates early-replication origins and replication timing.** **a**, Venn diagram showing the overlap between early-replication domains identified from our BrdU-IP-seq data ( $n = 3,362$ ) and Repli-seq data sets in ref. <sup>21</sup> ( $n = 4,727$ ). **b**, Distribution of ORC1 peaks and the 10-min BrdU signal (labelled in cells immediately after release from G1/S arrest) in length-normalized early-replication domains. **c**, Real-time PCR analysis of ChIP signals from H2A.Z, H4K20me2 and ORC1, and nascent-strand signals in siNC or siH2A.Z cells at target 3 (an early replication origin). **d**, Dot plot showing the

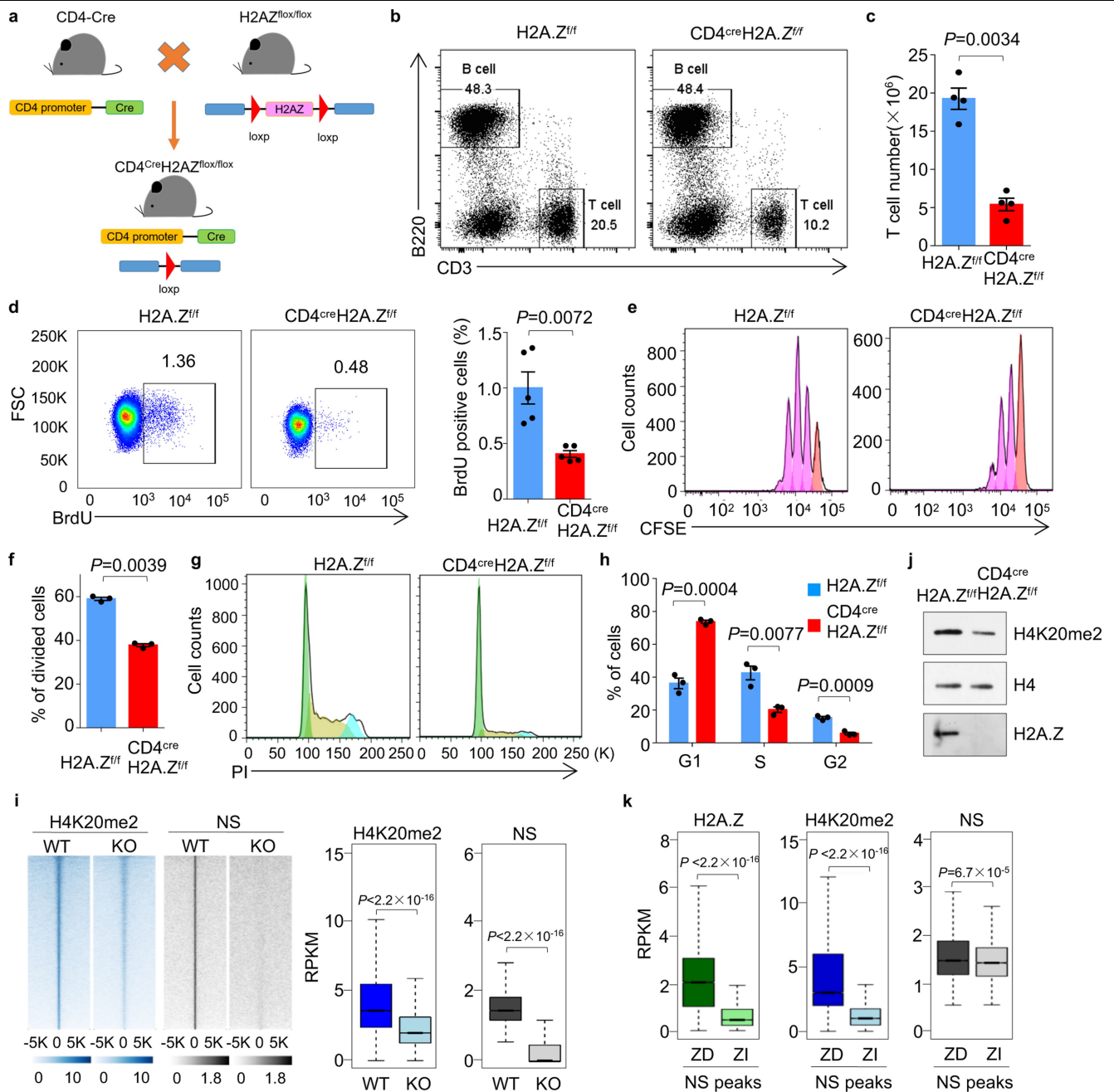
correlation between the replication timing of replication domains ( $n = 3,362$ ) identified by the 10-min BrdU signal and the dynamics of the 10-min BrdU signal after *H2AFZ* knockdown.  $r$ , Pearson's correlation coefficient. **e**, Genome tracks of an early-replication domain show the decreased 10-min BrdU signal after *H2AFZ* knockdown. RT, replication timing. Data in panel **c** are means  $\pm$  s.e.m.;  $n = 3$  biological replicates; two-tailed unpaired  $t$ -test. The BrdU-seq experiments in panels **b**, **e** were independently repeated twice with similar results.



**Extended Data Fig. 7 | H2A.Z regulates early-replication origins through SUV420H1. a**, FACS analysis of cell-cycle progression for siNC and siH2A.Z cells released from G1/S arrest. **b**, Statistical results from panel **a**. To quantify cell-cycle progression, we normalized the peak DAPI signal of all time points to the peak DAPI signal at 0 h (G1/S arrest). We defined the DAPI signal at the G1/S boundary as value 1; when cells entered G2/M phase, the value is near 2. **c**, Western blot analysis of H4K20me levels in *SUV420H1*-knockdown cells rescued by wild-type or mutant (R257A, K333A) *SUV420H1*. **d**, H4K20me2, NS and BrdU levels at early origins (targets 1, 3) and late origins (target 2) in

*SUV420H1*-knockdown ('Sh') cells rescued by wild-type or mutant *SUV420H1*. Data in panels **b**, **d** are means  $\pm$  s.e.m.;  $n=4$  biological replicates in **b**;  $n=3$  biological replicates in **d**; two-tailed unpaired *t*-test. The FACS experiment in panel **a** was independently repeated four times with similar results. The western blots in panel **c** were independently repeated three times with similar results. H4 was used as a loading control and sample processing control in panel **c**. For gel source data, see Supplementary Fig. 1. For the FACS gating strategy, see Supplementary Fig. 3.
















**Extended Data Fig. 8 | H2A.Z is essential for DNA replication and cell proliferation during T-cell activation.** **a**, Diagram showing the construction of mice with H2A.Z conditionally knocked out in T cells. **b**, FACS analysis of T cells and B cells from the spleen of wild-type (H2A.Z<sup>fl/fl</sup>) or H2A.Z CKO (CD4<sup>Cre</sup>H2A.Z<sup>fl/fl</sup>) mice. **c**, Statistical analysis of T-cell numbers from panel **b**. **d**, Left, FACS analysis of BrdU incorporation in CD8<sup>+</sup> T cells from the lymph node of wild-type and H2A.Z CKO mixed bone-marrow chimaeric mice. Right, statistical analysis of the percentage of BrdU-incorporating cells. FSC, forward scatter. **e**, FACS analysis of CFSE dilution in wild-type and H2A.Z CKO CD4<sup>+</sup> T cells upon anti-CD3 and anti-CD28 stimulation for 72 h. **f**, Statistical analysis of the percentage of divided cells from panel **e**. **g**, Cell-cycle analysis of wild-type and H2A.Z CKO CD4<sup>+</sup> T cells upon anti-CD3 and anti-CD28 stimulation for 48 h. PI, propidium iodide. **h**, Statistical analysis of cell-cycle distribution from panel **g**. **i**, Heat map and box plots showing the H4K20me2 and NS signals in 10-kb regions around H4K20me2 peaks ( $n=53,788$ ) and NS peaks ( $n=26,283$ ). **j**,

Western blots showing H4K20me2 levels in active T cells from wild-type and H2A.Z CKO mice. **k**, Box plots showing H2A.Z, H4K20me2 and NS signals in ZD-NS ( $n=18,382$ ) and ZI-NS ( $n=7,901$ ) regions. Data in panels **c**, **d**, **f** and **h** are means  $\pm$  s.e.m.;  $n=4$  biological replicates in panel **c**;  $n=5$  biological replicates in panel **d**;  $n=3$  biological replicates in panels **f**, **h**; two-tailed unpaired *t*-test. The data in panels **i**, **k** were analysed by two-tailed paired *t*-test and two-tailed unpaired *t*-test, respectively. FACS analyses in panels **b**, **d**, **e**, **g** were independently repeated four, five, three and three times, respectively, with similar results. The western blots in panel **j** were independently repeated three times with similar results, and H4 was used as a loading control and sample processing control. For the box plots in panels **i**, **k**, the centre lines represent medians, the box limits are the 25th and 75th percentiles, and the whiskers are the minimum to maximum values. For gel source data, see Supplementary Fig. 1. For FACS gating strategy, see Supplementary Fig. 3.

Extended Data Table 1 | Enriched DNA motifs in H2A.Z-regulated replication origins

Motif ID	<i>P</i> value	Sites	Logo
motif 1	$2.3 \times 10^{-36}$	365	
motif 2	$2.3 \times 10^{-36}$	342	
motif 3	$2.4 \times 10^{-34}$	1529	
motif 4	$2.1 \times 10^{-16}$	1528	
motif 5	$1 \times 10^{-15}$	1410	
motif 6	$1.2 \times 10^{-13}$	2017	
motif 7	$1.7 \times 10^{-13}$	1922	
motif 8	$1.1 \times 10^{-21}$	1535	
motif 9	$5.4 \times 10^{-20}$	3651	
motif 10	$2.2 \times 10^{-3}$	1712	
motif 11	$5.2 \times 10^{-6}$	1387	

The sequences of 400-bp genomic regions around the 20,978 plus-ZD-ORC1 peak summits were subjected to regulatory sequence analysis tools (RSATs) with default parameters.

## Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- | n/a                                 | Confirmed  |
|-------------------------------------|--|
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> The exact sample size ( <i>n</i> ) for each experimental group/condition, given as a discrete number and unit of measurement   |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided<br><i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A description of all covariates tested  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> For null hypothesis testing, the test statistic (e.g. <i>F</i> , <i>t</i> , <i>r</i> ) with confidence intervals, effect sizes, degrees of freedom and <i>P</i> value noted<br><i>Give P values as exact values whenever suitable.</i>                     |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings  |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> Estimates of effect sizes (e.g. Cohen's <i>d</i> , Pearson's <i>r</i> ), indicating how they were calculated   |

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

### Software and code

Policy information about [availability of computer code](#)

Data collection	Xcaliber 4.1.31.9; OLYMPUS FV1000 Ver.1.7a, Illumina Hiseq2000/Nova seq , BD FACS Diva 8.0.1
Data analysis	Proteome discoverer software 1.4, PEAKS 8.5, Launchpad 2.4, FV Viewer 2.0 , GraphPad Prism 6, FlowJo v.10 , Bowtie v2.2.5, bedtools v2.17.0, R v3.4.3, MACS v1.4.1, python v2.7.6, tophat v2.2.1, cufflinks v2.2.1, samtools v1.2.1, deeptools v2.3.5, IGV v2.3, PeakSeq v1.3, FASTX-Tools v 0.0.13

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

### Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

ChIP-seq, NS-seq, BrdU-seq and RNA-seq data have been deposited in Gene Expression Omnibus (GEO) under accession number GSE134988. Fig. 1-2 and Extended Data Fig.1-5, 7-8 have associated raw data in Supplementary Fig. 1.

# Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical methods were used to predetermine sample sizes. Sample sizes were determined based on previous experience.
Data exclusions	No data was excluded from the analysis.
Replication	All experiments were reliably reproduced. Each experiment was performed independently at least two times, but usually many more times.
Randomization	Mice were paired based on gender and age.
Blinding	Blinding was not performed due to the unambiguous nature of measurements and systematic analyses used in these experiments.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input type="checkbox"/>	<input checked="" type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data

### Methods

n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> ChIP-seq
<input type="checkbox"/>	<input checked="" type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

## Antibodies

### Antibodies used

For WB, ChIP, ChIP-seq:  
 Anti-H2A.Z, Abcam, ab4174, for WB , ChIP, ChIP-seq;  
 Anti-H4, Millipore, #05-858, for WB;  
 Anti-Flag, Sigma, s7452, for WB;  
 Anti-ORC1, Abcam, ab60, for WB;  
 Anti-ORC2, Abcam, ab68348 for WB;  
 Anti-MCM2, Abcam, ab4461, for WB;  
 Anti-RPA, Abcam, ab79398, for WB;  
 Anti-H4K20me1, Abcam, ab9051, for WB and ChIP-seq;  
 Anti-H4K20me2, Abcam, ab9052, for WB, ChIP and ChIP-seq;  
 Anti-H4K20me3, Abcam, ab9053, for WB and ChIP-seq;  
 Anti-Suv420H1, Novus, NBP1-78303, for WB;  
 Anti-tublin, Abcam, ab6046, for WB;  
 Anti-His, Invitrogen, MA1-21315, for WB;  
 Anti-H3, Cell signaling, #9715, for WB.

For flow cytometry  
 1. For Surface staining of mouse splenocytes  
 CD3 FITC (Clone: 145-2C11; Cat. 11-0031-82; Lot:4338511; eBioscience)  
 B220 PE (Clone: RA3-6B2; Cat. 12-0452; Lot: E01249-1634; eBioscience)  
 Anti-Brdu FITC (Clone: B44; Cat. 347583; Lot: 6090520; BD)  
 Live/Dead (Cat. L34959; Lot: 1921586; invitrogen)

2. For T cell Isolation and the analysis of proliferation and cell cycle  
 CD4 PerCp-Cy5.5 (Clone: RM4-5; Cat. 45-0042; Lot: 4304295; eBioscience)  
 CD8 APC (Clone: 53-6.7; Cat. 17-0081; Lot: E07057-1635; eBioscience)

CD25 PE (Clone: PC61.5; Cat. 12-0251; Lot: 4277529; eBioscience)  
 CD44 APC/Cy7 (Clone: IM7; Cat. 103028; Lot: B262797; Biolegend)  
 CD62L AF700 (Clone: MEL-14; Cat. 56-0621-821; Lot: 4306534; eBioscience)  
 Anti-mouse CD3e (Clone: 145-2C11; Cat. 100314 ; Lot:B267827; Biolegend)  
 Anti-mouse CD28 (Clone: 37.51; Cat. 102112; Lot:B261983; Biolegend)

## Validation

all antibodies used are commonly used in the field and have been validated in previous publications/by the manufacturer. References and manufacturer validations can be found here:  
 Anti-H2A.Z (Abcam ab4174): <https://www.abcam.com/histone-h2az-antibody-chip-grade-ab4174.html#top-496>  
 Anti-H4 (Millipore #05858): [https://www.merckmillipore.com/CN/zh/product/Anti-Histone-H4-Antibody-pan-clone-62-141-13-rabbit-monoclonal,MM\\_NF-05-858?ReferrerURL=https%3A%2F%2Fwww.google.com%2F&bd=1](https://www.merckmillipore.com/CN/zh/product/Anti-Histone-H4-Antibody-pan-clone-62-141-13-rabbit-monoclonal,MM_NF-05-858?ReferrerURL=https%3A%2F%2Fwww.google.com%2F&bd=1)  
 Anti-H3 (Cell signaling #9715): <https://www.cellsignal.com/products/primary-antibodies/histone-h3-antibody/9715>  
 Anti-H4K20me1 (Abcam ab9051): <https://www.abcam.com/histone-h4-mono-methyl-k20-antibody-chip-grade-ab9051.html>  
 Anti-H4K20me2 (Abcam ab9052): <https://www.abcam.com/histone-h4-di-methyl-k20-antibody-chip-grade-ab9052.html>  
 Anti-H4K20me3 (Abcam ab9053): <https://www.abcam.com/Histone-H4-tri-methyl-K20-antibody-ChIP-Grade-ab9053/reviews/45378>  
 Anti-H3K36me3 (Abcam Ab9050): <https://www.abcam.com/Histone-H3-tri-methyl-K36-antibody-ChIP-Grade-ab9050/reviews/61366>  
 Anti-ORC1 (Abcam ab60): <https://www.abcam.com/orc1-antibody-7f61-chip-grade-ab60.html>  
 Anti-PCNA (Abcam ab29): <https://www.abcam.com/pcna-antibody-pc10-ab29.html>  
 Anti-BrdU (BD Biosciences BD44): <https://www.bdbiosciences.com/us/applications/research/apoptosis/purified-antibodies/purified-mouse-anti-brdu-b44/p/347580>  
 Anti-Flag (Sigma F7452): <https://www.sigmaaldrich.com/catalog/product/sigma/f7452?lang=zh&region=CN>  
 Anti-ORC2 (Abcam ab68348): <https://www.abcam.com/orc2-antibody-ab68348.html>  
 Anti-MCM2 (Abcam ab4461): <https://www.abcam.com/mcm2-antibody-ab4461.html>  
 Anti-RPA (Abcam ab79398): <https://www.abcam.com/rpa70-antibody-epr3472-ab79398.html>  
 Anti-Suv420H1 (Novus, NBP1-78303): [https://www.novusbio.com/products/suv420h1-antibody\\_nbp1-78303](https://www.novusbio.com/products/suv420h1-antibody_nbp1-78303)  
 Anti-Tubulin (Abcam Ab6046): <https://www.abcam.com/beta-tubulin-antibody-loading-control-ab6046.html>  
 Anti-His (Invitrogen MA1-21315): [https://www.fishersci.com/shop/products/anti-6x-his-epitope-tag-clone-his-h8/MA121315?gclid=Cj0KCQjwoKzBRCSARIsAITcwXEHNjo-RWL5Ce9YPTuOXqXAPKmQjn-y3MVTyALYyMfjCpWgli4DGwkaAmT7EALw\\_wcB&ef\\_id=Cj0KCQjwoKzBRCSARIsAITcwXEHNjo-RWL5Ce9YPTuOXqXAPKmQjn-y3MVTyALYyMfjCpWgli4DGwkaAmT7EALw\\_wcB:s&cid=SEM\\_GAW\\_20190909\\_CB88IN&ppc\\_id=FisherSciNonbrand\\_goog\\_6492004464\\_80722830231\\_AntibodiesDSA\\_b\\_381598720312\\_791621048115855394&s\\_kwcid=AL14428!3!381598720312!b!!g!!](https://www.fishersci.com/shop/products/anti-6x-his-epitope-tag-clone-his-h8/MA121315?gclid=Cj0KCQjwoKzBRCSARIsAITcwXEHNjo-RWL5Ce9YPTuOXqXAPKmQjn-y3MVTyALYyMfjCpWgli4DGwkaAmT7EALw_wcB&ef_id=Cj0KCQjwoKzBRCSARIsAITcwXEHNjo-RWL5Ce9YPTuOXqXAPKmQjn-y3MVTyALYyMfjCpWgli4DGwkaAmT7EALw_wcB:s&cid=SEM_GAW_20190909_CB88IN&ppc_id=FisherSciNonbrand_goog_6492004464_80722830231_AntibodiesDSA_b_381598720312_791621048115855394&s_kwcid=AL14428!3!381598720312!b!!g!!)  
 Anti-CD3-FITC (eBioscience 11-0031-82 clone 145-2C11): <https://www.thermofisher.com/cn/zh/antibody/product/CD3e-Antibody-clone-145-2C11-Monoclonal/11-0031-82>  
 Anti-CD4-PerCp-Cy5.5 (eBioscience 45-0042 clone RM4-5): <https://www.thermofisher.com/cn/zh/antibody/product/CD4-Antibody-clone-RM4-5-Monoclonal/35-0042-82>  
 Anti-CD8-APC (Biolegend 17-0081 clone 53-6.7): <https://www.thermofisher.com/cn/zh/antibody/product/CD8-alpha-Antibody-clone-5H10-Monoclonal/MCD0805>  
 Anti-CD44-APC/Cy7 (Biolegend 103028): <https://www.biolegend.com/en-us/products/apc-cy7-anti-mouse-human-cd44-antibody-3933>  
 Anti-CD62L-AF700 (eBioscience 56-0621-821 Clone MEL-14): <https://www.thermofisher.com/cn/zh/antibody/product/CD62L-L-Selectin-Antibody-clone-MEL-14-Monoclonal/56-0621-82>  
 Anti-B220-PE (eBioscience 12-0452 clone RA3-6B2): <https://www.thermofisher.com/antibody/product/12-0452-85.html?CID=AFLCA-12-0452-85>  
 Anti-CD25-PE (eBioscience 12-0251-82): <https://www.thermofisher.com/cn/zh/antibody/product/CD25-Antibody-clone-PC61-5-Monoclonal/12-0251-82>  
 Live/Dead (Invitrogen L34959): <https://www.thermofisher.com/order/catalog/product/L34959>  
 Anti-mouse CD28 (Biolegend 102112): <https://www.biolegend.com/en-us/products/purified-anti-mouse-cd28-antibody-117>  
 Anti-mouse CD3 (Biolegend 100314): <https://www.labome.com/product/BioLegend/100314.html>

## Eukaryotic cell lines

### Policy information about cell lines

Cell line source(s)	HeLa cell line was originally obtained from ATCC.
Authentication	Identity of HeLa cell line was frequently checked by the morphological features, but not authenticated.
Mycoplasma contamination	The cell lines were regularly tested for mycoplasma contamination. And if have contamination, we immediately treated cell lines with plasmocin(Invivogen).
Commonly misidentified lines (See <a href="#">ICLAC</a> register)	No commonly misidentified cell lines were used.

## Animals and other organisms

### Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	H2A.Z flox mice were purchased from Riken Bioresource Center(RBRC05765). CD4-Cre transgenic mice were purchased from The Jackson Laboratory. H2A.Zflox/flox CD4-Cre mice were generated through crossing H2A.Zflox/flox mice to CD4-Cre transgenic mice. All mice were housed under specific pathogen-free conditions in the animal care facilities at the Institute of Biophysics, Chinese Academy of Sciences.
--------------------	--



All mice used for experiments were 6-10 weeks-old. Age and sex matched female or male mice were used for each experiment.

#### Wild animals

No wild animals were involved.

#### Field-collected samples

No field-collected samples were used.

#### Ethics oversight

All animal experiments were performed in accordance with the guidelines of the Institute of Biophysics, Chinese Academy of Sciences, using protocols approved by the Institutional Laboratory Animal Care and Use Committee.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## ChIP-seq

### Data deposition

- ☒ Confirm that both raw and final processed data have been deposited in a public database such as [GEO](https://www.ncbi.nlm.nih.gov/geo/).
- ☒ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

#### Data access links

May remain private before publication.

<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE134988>

#### Files in database submission

GSM3983183 hela-siNC-H2A.Z\_rep1  
 GSM3983184 hela-siNC-H2A.Z\_rep2  
 GSM3983185 hela-siH2A.Z-H2A.Z\_rep1  
 GSM3983186 hela-siH2A.Z-H2A.Z\_rep2  
 GSM3983187 hela-siNC-H4K20me1\_rep1  
 GSM3983188 hela-siNC-H4K20me1\_rep2  
 GSM3983189 hela-siNC-H4K20me2\_rep1  
 GSM3983190 hela-siNC-H4K20me2\_rep2  
 GSM3983191 hela-siNC-H4K20me3\_rep1  
 GSM3983192 hela-siNC-H4K20me3\_rep2  
 GSM3983193 hela-siNC-native-input  
 GSM3983194 hela-siH2A.Z-native-input  
 GSM3983195 hela-siH2A.Z-H4K20me1\_rep1  
 GSM3983196 hela-siH2A.Z-H4K20me1\_rep2  
 GSM3983197 hela-siH2A.Z-H4K20me2\_rep1  
 GSM3983198 hela-siH2A.Z-H4K20me2\_rep2  
 GSM3983199 hela-siH2A.Z-H4K20me3\_rep1  
 GSM3983200 hela-siH2A.Z-H4K20me3\_rep2  
 GSM3983201 hela-siNC-ORC1\_rep1  
 GSM3983202 hela-siNC-ORC1\_rep2  
 GSM3983203 hela-siH2A.Z-ORC1\_rep1  
 GSM3983204 hela-siH2A.Z-ORC1\_rep2  
 GSM3983205 hela-siNC-NS\_rep1  
 GSM3983206 hela-siNC-NS\_rep2  
 GSM3983207 hela-siNC-Rnase-treated-NS\_rep1  
 GSM3983208 hela-siNC-Rnase-treated-NS\_rep2  
 GSM3983209 hela-siNC-genome  
 GSM3983210 hela-siH2A.Z-NS  
 GSM3983211 hela-siH2A.Z-Rnase-treated-NS  
 GSM3983212 hela-siH2A.Z-genome  
 GSM3983213 hela-siNC-BrdU-10min\_rep1  
 GSM3983214 hela-siNC-BrdU-10min\_rep2  
 GSM3983215 hela-siNC-BrdU-1h  
 GSM3983216 hela-siNC-BrdU-3h  
 GSM3983217 hela-siNC-BrdU-6h  
 GSM3983218 hela-siH2A.Z-10min\_rep1  
 GSM3983219 hela-siH2A.Z-10min\_rep2  
 GSM3983220 hela-siH2A.Z-1h  
 GSM3983221 hela-siH2A.Z-3h  
 GSM3983222 hela-siH2A.Z-6h  
 GSM3983223 hela-siNC-RNA-seq\_rep1  
 GSM3983224 hela-siNC-RNA-seq\_rep2  
 GSM3983225 hela-siH2A.Z-RNA-seq\_rep1  
 GSM3983226 hela-siH2A.Z-RNA-seq\_rep2  
 GSM3983227 mouse-active-T-cell-siNC-H4K20me2  
 GSM3983228 mouse-active-T-cell-siH2A.Z-H4K20me2  
 GSM3983229 mouse-active-T-cell-siNC-input  
 GSM3983230 mouse-active-T-cell-siH2A.Z-input  
 GSM3983231 mouse-active-T-cell-siNC-NS  
 GSM3983232 mouse-active-T-cell-siNC-Rnase-treated-NS  
 GSM3983233 mouse-active-T-cell-siH2A.Z-NS

GSM3983234 mouse-active-T-cell-siH2A.Z-Rnase-treated-NS  
GSM3993419 hela-WT-input

Genome browser session  
(e.g. [UCSC](#))

Not used

## Methodology

Replicates

BrdU, ORC1 and Histone ChIP-seq experiments were performed in duplicates with an input control for each experiment. Nascent strand (NS) seq experiments were performed in duplicates with an input control for each experiment.

Sequencing depth

hela-WT-H2A.Z\_rep1, single-end, read number: 29217798, read length: 50 bp  
hela-WT-H2A.Z\_rep2, single-end, read number: 36786000, read length: 50 bp  
hela-H2A.Z-KD-H2A.Z\_rep1, single-end, read number: 30506757, read length: 50 bp  
hela-H2A.Z-KD-H2A.Z\_rep2, single-end, read number: 30327697, read length: 50 bp  
hela-WT-H4K20me1\_rep1, single-end, read number: 27007998, read length: 50 bp  
hela-WT-H4K20me2\_rep1, single-end, read number: 46342466, read length: 50 bp  
hela-WT-H4K20me2\_rep2, single-end, read number: 21530865, read length: 50 bp  
hela-WT-H4K20me3\_rep1, single-end, read number: 35063713, read length: 50 bp  
hela-WT-input\_rep1, single-end, read number: 35063713, read length: 50 bp  
hela-WT-input\_rep2, single-end, read number: 30813565, read length: 50 bp  
hela-H2A.Z-KD-input\_rep1, single-end, read number: 33781291, read length: 50 bp  
hela-H2A.Z-KD-input\_rep2, single-end, read number: 31197655, read length: 50 bp  
hela-H2A.Z-KD-H4K20me1\_rep1, single-end, read number: 33187537, read length: 50 bp  
hela-H2A.Z-KD-H4K20me2\_rep1, single-end, read number: 41079876, read length: 50 bp  
hela-H2A.Z-KD-H4K20me2\_rep2, single-end, read number: 22572525, read length: 50 bp  
hela-H2A.Z-KD-H4K20me3\_rep1, single-end, read number: 32816918, read length: 50 bp  
hela-WT-ORC1\_rep1, single-end, read number: 26780831, read length: 50 bp  
hela-WT-ORC1\_rep2, single-end, read number: 27504126, read length: 50 bp  
hela-H2A.Z-KD-ORC1\_rep1, single-end, read number: 39982452, read length: 50 bp  
hela-H2A.Z-KD-ORC1\_rep2, single-end, read number: 21389952, read length: 50 bp  
hela-WT-NS\_rep1, single-end, read number: 22799872, read length: 50 bp  
hela-WT-NS\_rep2, single-end, read number: 24639890, read length: 50 bp  
hela-WT-Rnase-treated-NS\_rep1, single-end, read number: 18786498, read length: 50 bp  
hela-WT-Rnase-treated-NS\_rep2, single-end, read number: 53131336, read length: 51 bp  
hela-WT-genome-NS\_rep1, single-end, read number: 16940577, read length: 50 bp  
hela-WT-genome-NS\_rep2, single-end, read number: 36229449, read length: 50 bp  
hela-H2A.Z-KD-NS\_rep1, single-end, read number: 33595669, read length: 50 bp  
hela-H2A.Z-KD-Rnase-treated-NS\_rep1, single-end, read number: 23120958, read length: 50 bp  
hela-H2A.Z-KD-genome-NS\_rep1, single-end, read number: 25083887, read length: 50 bp  
hela-WT-BrdU-10min\_rep1, single-end, read number: 35791042, read length: 50 bp  
hela-WT-BrdU-10min\_rep2, single-end, read number: 23413789, read length: 50 bp  
hela-WT-BrdU-1h\_rep1, single-end, read number: 42484612, read length: 50 bp  
hela-WT-BrdU-3h\_rep1, single-end, read number: 33748760, read length: 50 bp  
hela-WT-BrdU-6h\_rep1, single-end, read number: 21086656, read length: 50 bp  
hela-H2A.Z-KD-10min\_rep1, single-end, read number: 41527806, read length: 50 bp  
hela-H2A.Z-KD-10min\_rep2, single-end, read number: 24074710, read length: 50 bp  
hela-H2A.Z-KD-1h\_rep1, single-end, read number: 47490348, read length: 50 bp  
hela-H2A.Z-KD-3h\_rep1, single-end, read number: 25062487, read length: 50 bp  
hela-H2A.Z-KD-6h\_rep1, single-end, read number: 19341066, read length: 50 bp  
hela-WT-RNA-seq, pair-end, read number: 41904139, read length: 150 bp  
hela-H2A.Z-KD-RNA-seq, pair-end, read number: 44240416, read length: 150 bp  
mouse-active-T-cell-WT-H4K20me2, single-end, read number: 28935057, read length: 50 bp  
mouse-active-T-cell-H2A.Z-cKO-H4K20me2, single-end, read number: 24039043, read length: 50 bp  
mouse-active-T-cell-WT-input, single-end, read number: 24703906, read length: 50 bp  
mouse-active-T-cell-H2A.Z-cKO-input, single-end, read number: 28063843, read length: 50 bp  
mouse-active-T-cell-WT-NS, pair-end, read number: 41513960, read length: 150 bp  
mouse-active-T-cell-H2A.Z-cKO-NS, pair-end, read number: 47561656, read length: 150 bp  
mouse-active-T-cell-WT-NS, pair-end, read number: 62444070, read length: 150 bp  
mouse-active-T-cell-H2A.Z-cKO-NS, pair-end, read number: 57631120, read length: 150 bp

Antibodies

Anti-H2A.Z, Abcam, ab4174; Anti-H4K20me1, Abcam, ab9051; Anti-H4K20me2, Abcam, ab9052; Anti-H4K20me3, Abcam, ab9053; Anti-BrdU, BD, BD44. The ORC1 chip used the M-280 streptavidin dynabead.

Peak calling parameters

Reads were uniquely mapped to genome using bowtie2. Peaks were called using MACS (--shiftsize=75) and PeakSeq (fdr below 0.5%)

Data quality

Read with high quality were retained using Fastx\_toolkit, unique reads were used for peak calling with FDR below 0.5%

Software

Bowtie v2.2.5, bedtools v2.17.0, R v3.4.3, MACS v1.4.1, python v2.7.6, tophat v2.2.1, cufflinks v2.2.1, samtools v1.2.1, deeptools v2.3.5, IGV v2.3, PeakSeq v1.3, FASTX-Tools v 0.0.13,

# Flow Cytometry

## Plots

Confirm that:

- ☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- ☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- ☒ All plots are contour plots with outliers or pseudocolor plots.
- ☒ A numerical value for number of cells or percentage (with statistics) is provided.

## Methodology

Sample preparation

Spleens and lymph node cells from mice were harvested and homogenized. Cells were suspended in FACS buffer(2% FBS in PBS) and cell suspensions were filtered through 70um cell strainer. The suspension was washed with cold FACS buffer and centrifuged at 500g for 5min.

Instrument

LSRFortessa, FACS Aria III

Software

BD FACS Diva 8.0.1 software was used for data collection and FlowJo v.10 was used for data analysis.

Cell population abundance

For the spleen and lymph node cells, there are~25% of CD4+ T cell, ~20% of CD8+ T cell in single alive lymphocytes. For CD4+ T cells, there are ~85% of CD62L+CD44- naïve T cells.  
For the spleen lymphocytes, there are~20% of CD3+ T cell and ~50% of B220+ B cell in WT mouse, ~10% of CD3+ T cell in CKO mouse.  
For flow sorting, post-sort cells were analyzed on BD Aria III and the purity was at least 95%.  
For the cell cycle, there are ~65% of G1/G0 phase cell, ~25% of S phase cell and ~10% of M phase cell in WT group, ~80% of G1/G0 phase cell, ~10% of S phase cell and ~10% of M phase cell in H2A.Z KO group.

For flow sorting, post-sort cells were analyzed on BD Aria III and the purity was at least 95%.

Gating strategy

I. For T cell analysis and sorting  
I-a. For spleen lymphocytes, T cells were gated with CD3 and B cell gated with B220.  
I-b. For sorting naïve T cells, single alive lymphocytes were first gated with CD4 and CD8, then naïve T cells were gated with CD44 and CD62L, CD44 low and CD62L high cells were resident naïve T cells.  
II. For CFSE analysis  
Divided cells: single alive lymphocytes were gated with CFSE+, and CFSE diluted cells are divided cells.  
III. For cell cycle analysis  
Single alive lymphocytes were gated with Brdu and PI. Brdu low and PI low cells were resident G1/G0 phase cells; Brdu high and PI middle cells were resident S phase cell cells; Brdu low and PI high cells were resident M phase cell cells.

- ☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

# Author Correction: Weak average liquid- cloud-water response to anthropogenic aerosols

<https://doi.org/10.1038/s41586-019-1838-3>

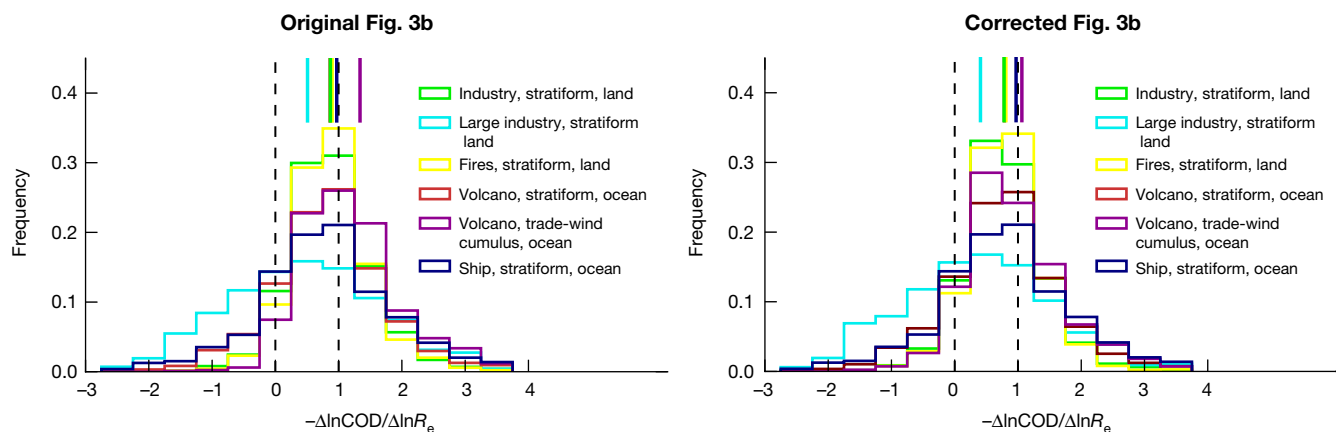
Correction to: *Nature* <https://doi.org/10.1038/s41586-019-1423-9>

Published online 31 July 2019

Velle Toll, Matthew Christensen, Johannes Quaas & Nicolas Bellouin

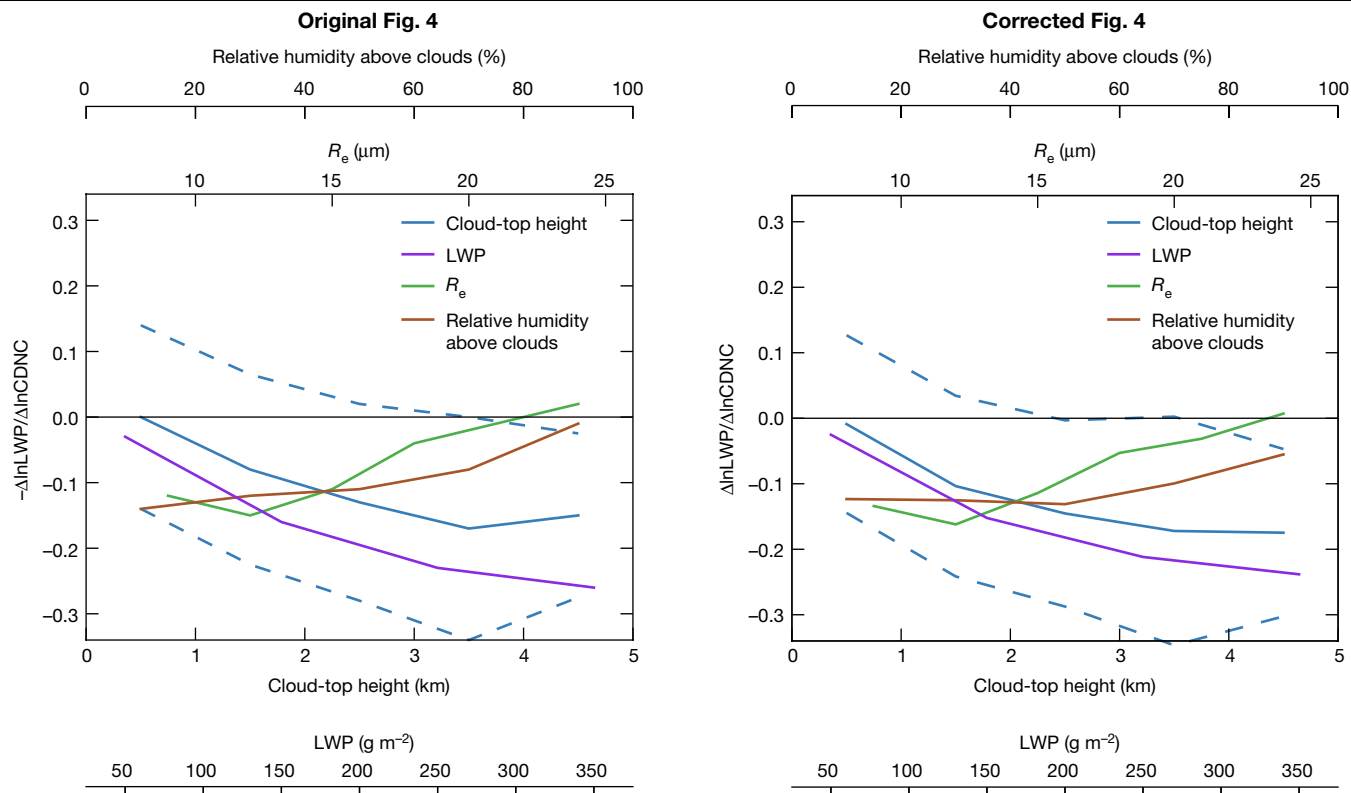
In this Article, a coding mistake occurred when calculating  $-\frac{\Delta \ln \text{COD}}{\Delta \ln R_e}$  and  $\frac{\Delta \ln \text{LWP}}{\Delta \ln \text{CDNC}}$ , in which COD denotes the cloud optical depth;  $R_e$  denotes the cloud droplet effective radius; LWP denotes the liquid water path; and CDNC denotes the cloud droplet number concentration. The natural logarithm of those cloud properties was mistakenly taken before, instead of after, calculating the track segment average. The mistake has been corrected in Figs. 3b, 4, 6 and Extended Data Tables 2, 4 of the original Article, and the incorrect, published figures and tables are shown as Figs. 1–5 of this Amendment, for transparency to readers. The conclusions of the paper, including the main conclusion that changes in cloud water caused by aerosols exert a weak climate-warming effect, and all other figures and tables, are not affected.

The correction has minor effects on the results of the paper, which remain qualitatively the same. The radiative forcing exerted by LWP adjustments is now estimated at  $+0.15 \text{ W m}^{-2}$ , instead of  $+0.12 \text{ W m}^{-2}$ . Consequently, the observed decrease in cloud water now offsets 29%, up from 23%, of the global climate-cooling effect caused by aerosol-induced increases in the concentration of cloud droplets. In addition, the y-axis label in Fig. 4 should be  $\Delta \ln \text{LWP} / \Delta \ln \text{CDNC}$  without a minus sign. The original Article has been corrected online.

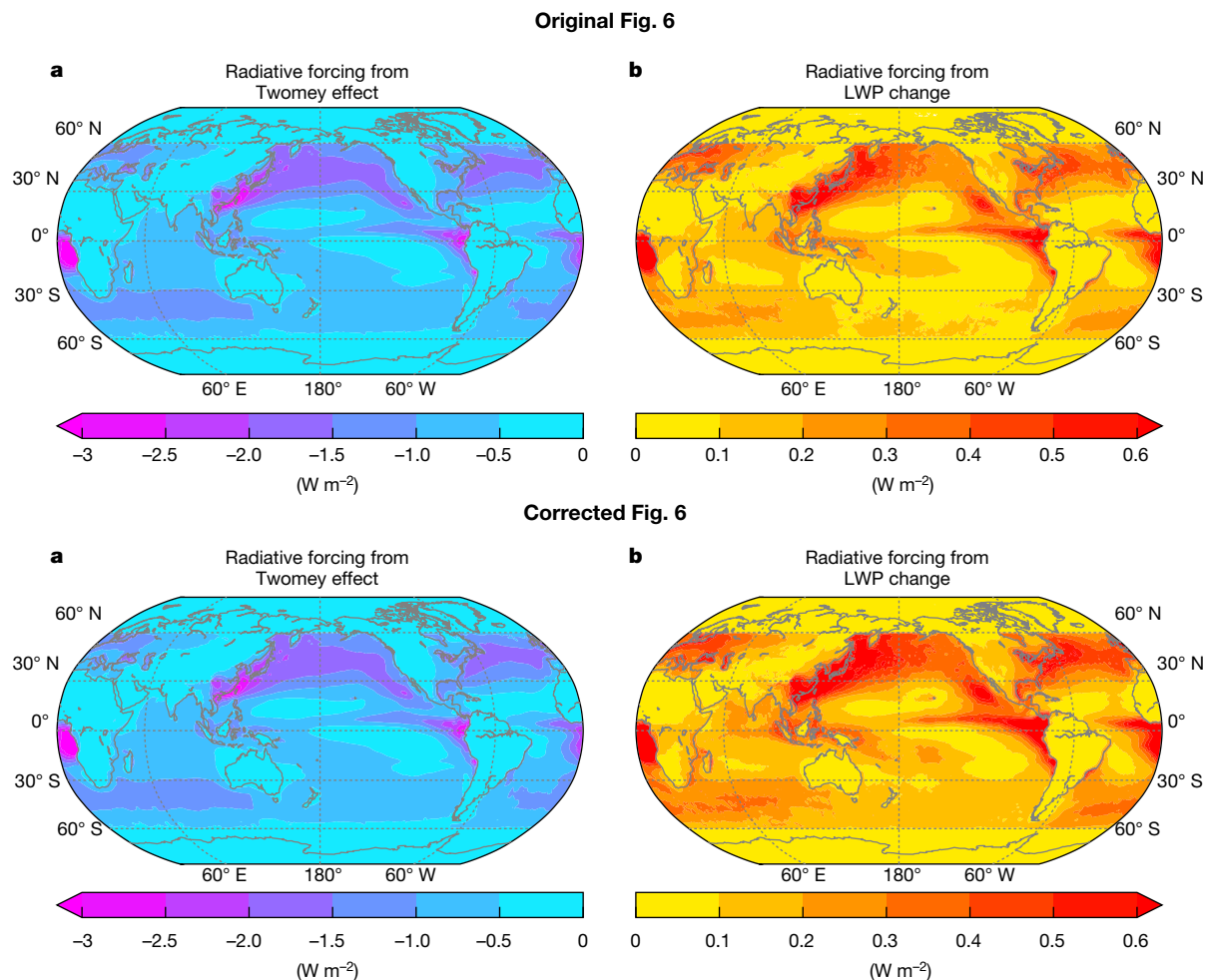


**Fig. 1** | This figure displays the corrected and the incorrect published Fig. 3b of the original Article.

## Corrections & amendments



**Fig. 2** | This figure displays the corrected and the incorrect published Fig. 4 of the original Article.



**Fig. 3** | This figure displays the corrected and the incorrect published Fig. 6 of the original Article.



Original Extended Data Table 2				
Type of track	Number of track observations	Cloud top height [m]	Relative humidity [%]	$\Delta \ln \text{LWP} / \Delta \ln \text{CDNC}$
Industry St land	6478	2259	42.6	-0.071 (0.003)
Large industry St land	984	2862	34.2	-0.215 (0.018)
Fires St land	1389	2532	46.5	-0.088 (0.006)
Volcano St ocean	2521	1311	39.7	-0.107 (0.007)
Volcano Cu ocean	831	1471	76.1	+0.051 (0.012)
Ship St ocean	793	721	26.6	-0.021 (0.013)

Corrected Extended Data Table 2				
Type of track	Number of track observations	Cloud top height [m]	Relative humidity [%]	$\Delta \ln \text{LWP} / \Delta \ln \text{CDNC}$
Industry St land	6478	2259	42.6	-0.098 (0.003)
Large industry St land	984	2862	34.2	-0.249 (0.017)
Fires St land	1389	2532	46.5	-0.081 (0.006)
Volcano St ocean	2521	1311	39.7	-0.109 (0.007)
Volcano Cu ocean	831	1471	76.1	-0.021 (0.012)
Ship St ocean	793	721	26.6	-0.028 (0.013)

**Fig. 4 |** This figure displays the corrected and the incorrect published Extended Data Table 2 of the original Article.

Original Extended Data Table 4		
$R_e$ interval [ $\mu\text{m}$ ]	$\Delta \ln \text{LWP} / \Delta \ln \text{CDNC}$ volcano and ship tracks over ocean	$\Delta \ln \text{LWP} / \Delta \ln \text{CDNC}$ industry and fire tracks over land
$R_e < 10.5$	-0.115 (0.013)	-0.124 (0.014)
$10.5 \leq R_e < 13.5$	-0.140 (0.012)	-0.145 (0.005)
$13.5 \leq R_e < 16.5$	-0.068 (0.015)	-0.104 (0.006)
$16.5 \leq R_e < 19.5$	-0.004 (0.017)	-0.043 (0.006)
$19.5 \leq R_e < 22.5$	+0.052 (0.033)	-0.035 (0.009)
$R_e \geq 22.5$	+0.106 (0.035)	-0.008 (0.017)

Corrected Extended Data Table 4		
$R_e$ interval [ $\mu\text{m}$ ]	$\Delta \ln \text{LWP} / \Delta \ln \text{CDNC}$ volcano and ship tracks over ocean	$\Delta \ln \text{LWP} / \Delta \ln \text{CDNC}$ industry and fire tracks over land
$R_e < 10.5$	-0.145 (0.012)	-0.129 (0.009)
$10.5 \leq R_e < 13.5$	-0.165 (0.010)	-0.161 (0.005)
$13.5 \leq R_e < 16.5$	-0.094 (0.012)	-0.122 (0.006)
$16.5 \leq R_e < 19.5$	-0.026 (0.016)	-0.062 (0.006)
$19.5 \leq R_e < 22.5$	+0.056 (0.022)	-0.058 (0.009)
$R_e \geq 22.5$	+0.116 (0.030)	-0.026 (0.011)

**Fig. 5 |** This figure displays the corrected and the incorrect published Extended Data Table 4 of the original Article.

# Author Correction: Molecular architecture of lineage allocation and tissue organization in early mouse embryo

---

<https://doi.org/10.1038/s41586-019-1887-7>

---

Correction to: *Nature* <https://doi.org/10.1038/s41586-019-1469-8>

---

Published online 7 August 2019

---

Guangdun Peng, Shengbao Suo, Guizhong Cui, Fang Yu, Ran Wang,  
Jun Chen, Shirui Chen, Zhiwen Liu, Guoyu Chen, Yun Qian,  
Patrick P. L. Tam, Jing-Dong J. Han & Naihe Jing

---

In Extended Data Fig. 6a of this Letter, for consistency with the main figures, the labels 'A' and 'P' should be 'Epi1' and 'Epi2', respectively, and labels 'EA' and 'EP' should be 'En1' and 'En2', respectively. In Supplementary Table 6, the content of the table was wrongly calculated owing to an error during coding. In Supplementary Table 7, the label 'Epi1.E7.5' should be 'Ect1.E7.5' for consistency with the main figures. These errors do not affect the conclusions of the Letter. All errors have been corrected online.

# **Publisher Correction: In vivo imaging of mitochondrial membrane potential in non-small-cell lung cancer**

---

<https://doi.org/10.1038/s41586-019-1890-z>

---

Correction to: *Nature* <https://doi.org/10.1038/s41586-019-1715-0>

---

Published online 30 October 2019

---

**Milica Momcilovic, Anthony Jones, Sean T. Bailey,  
Christopher M. Waldmann, Rui Li, Jason T. Lee, Gihad Abdelhady,  
Adrian Gomez, Travis Holloway, Ernst Schmid, David Stout,  
Michael C. Fishbein, Linsey Stiles, Deepa V. Dabir,  
Steven M. Dubinett, Heather Christofk, Orian Shirihai,  
Carla M. Koehler, Saman Sadeghi & David B. Shackelford**

---

In this Article, owing to an error during the production process, author Jason T. Lee was erroneously associated with affiliation 3 (The Mouse Phase I Unit, Lineberger School of Medicine at the University of North Carolina Chapel Hill, Chapel Hill, NC, USA), and should have instead been associated with affiliations 2, 4 and 5 (Department of Molecular and Medical Pharmacology, David Geffen School of Medicine at the University of California, Los Angeles, CA, USA; Crump Institute for Molecular Imaging, David Geffen School of Medicine at the University of California, Los Angeles, CA, USA; and Jonsson Comprehensive Cancer Center, David Geffen School of Medicine at the University of California, Los Angeles, CA, USA). This has been corrected online.



TONY WU/NATURE PICTURE LIBRARY

The fate of whales provided a hook for helping policymakers to understand how science works in practice.

## POLICY TALES AND THE SECRET LIFE OF WHALES

Palaeontologist finds a way to convey science to business leaders at the World Economic Forum. **By Nick Pyenson**

**A**s a palaeontologist who works with fossils of large, extinct ocean predators, I tend to think that the story of our future has already been written in the geological past. The same rocks that preserve the remains of ancient whales tell us about dramatic sea-level rises that might be matched in our future, if global warming continues. As we begin to encounter geological-scale global changes in our own lifetimes, the past of this planet is a guide to what might happen. It's hard for me to accept that scientists can explain how whale bones end up on

mountain tops but we can't find leadership to forestall glacial melting.

Leadership was definitely on my mind when I attended the World Economic Forum (WEF) Annual Meeting of the New Champions 2019 this past July in Dalian, China, to talk about

**"Many elected leaders pay little attention to scientific evidence."**

the secret life of whales to non-scientists from the business and policy fields. I thought the narrative of where whales originated, and how their fate today is inextricably linked with ours, would have traction at the WEF. I planned to use the fate of whales not just as a hook for amazing facts, but as a vehicle for understanding how science works in practice. I was unsure about how my presentation would land; after all, many elected leaders pay little attention to scientific evidence, often wilfully undermining it or happily ignoring it.

I knew that the WEF was important: much



VISUAL CHINA GROUP/GETTY

The Annual Meeting of the New Champions 2019, held by the World Economic Forum in Dalian, China.

of its influence comes not from the named attendees, but from using its platform and network to affect change across many areas of governance. I had also joined the WEF's Young Scientists community, which drew together a select group of early-career scientists from around the world for a two-year 'journey' (now three years). I was reassured by the other young scientists, who shared my hope for science at the WEF; their presence in the audience gave me much-needed support. Fortunately, my talk went down well.

So, what business do scientists have at a meeting such as the WEF? And what are the lessons for scientists who want to communicate their relevance and the overall importance of science to global leaders?

**Your expertise matters.** Scientists at the cutting edge of their fields have credibility that is hard-won and long-lasting. Use the opportunity granted by credibility to share information with people outside your normal scientific network.

**Scientific findings have value.** They don't necessarily show up in investor reports, but have ways of being durable and surprising. Scientists should speak about the value of scientific knowledge so that it isn't opaque

or discounted as irrelevant.

**Stories of discovery are exciting.** Whether it involves pandemics or neutrinos, don't underestimate the thrill of discoveries. Scientists are experts at pursuing knowledge, and we should speak clearly about how

**“Scientists should speak about the value of scientific knowledge.”**

we work things out. Part of the excitement is not always knowing the answers to our questions, together with the unexpected challenges and insights along the way. Told correctly, these testimonies can inspire and motivate a range of audiences for a long time.

**Facts need narrative.** It's clear that facts aren't always enough to capture interest or sway public opinion. Scientists can use the first-person narrative in unique ways. Combining subject expertise and storytelling savvy can give scientists influence in these multi-stakeholder meetings. The best presentations by scientists in Dalian did a lot more than merely translate jargon – the scientists used their subject knowledge and

the power of narrative to captivate and connect with their audience. Giving entertaining, engaging talks requires knowing the facts, but also recognizing what details to omit.

Scientists, of course, aren't great at everything. Although science has a part to play in nearly all of the 17 United Nations Sustainable Development Goals, scientists alone would have a hard time writing them. For scientists who want to step up to the multi-stakeholder table, they need to understand the priorities of political and business leaders – after all, we can't expect world leaders to become scientific experts in their spare time.

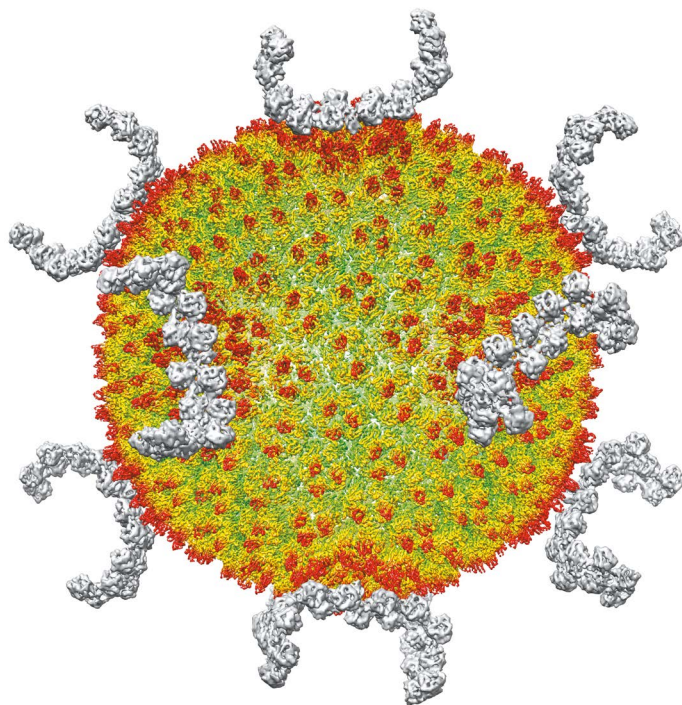
The big decisions of our time, including how we respond to future sea-level rise, need to be made by people who understand the complexity of the world, and who possess both confidence with creative problem-solving and the patience needed to play the long game. Scientists have these traits in abundance, along with the credibility and competence to make a difference at the table of global leadership, which the world certainly needs.

**Nick Pyenson** is a research geologist and curator of fossil marine mammals at the National Museum of Natural History, Smithsonian Institution, Washington DC, USA. He is the author of *Spying on Whales* (2018).



# TECHNOLOGIES TO WATCH IN 2020

Thought leaders describe the tech developments that could have a big impact in the coming year. **By Esther Landhuis**



The virus SH1, reconstructed from images obtained using cryogenic electron microscopy.

## HONGWEI WANG BETTER CRYO-EM SAMPLES

In two or three years, I think that transmission cryogenic electron microscopy (cryo-EM) will become the most powerful tool for deciphering the structures of macromolecules. These structures are crucial for understanding biochemical mechanisms and drug development, and methods for solving them more efficiently can speed up such work.

In cryo-EM, quickly freezing biological specimens in liquid nitrogen helps to preserve the molecules' water content and reduces damage from the high-energy electrons used for imaging. But specimen preparation is a major bottleneck: if you don't have a good specimen, you have nothing to image. Biological specimens often contain proteins, which unravel at the surface of the thin liquid layers used in the freezing process.

To prevent this unfolding, researchers are developing approaches that anchor proteins on to two-dimensional materials – such as the carbon lattice graphene – before

applying the liquid droplets. That way, they can make the droplets even smaller while keeping the protein away from the air–water interface<sup>1</sup>.

Some laboratories place nanolitre-sized samples directly on to a surface<sup>2</sup>, instead of using cumbersome older methods that draw excess liquid away from larger droplets. Other methods use a focused ion beam to slice frozen cells into layers thinner than 100 nanometres, allowing researchers to study molecules in their cellular contexts<sup>3</sup>.

Solving a molecular structure with cryo-EM typically requires collecting and analysing as many as 10,000 images, representing several weeks to a month of work. Many images are imperfect, so we have to discard them. But theoretically, a few dozen pictures should be enough, and it would take less than a day to collect and analyse them. This increased throughput could help us to understand disease mechanisms and develop drugs more efficiently.

**Hongwei Wang** is a structural biologist at Tsinghua University in Beijing.

## SARAH WOODSON IMPROVING RNA ANALYSIS

I'm keeping my eye on long-read RNA sequencing and live-cell imaging using light-up RNA strands called aptamers. These technologies are still maturing, but I expect big changes in the next year or two.

Short-read sequencing has changed the field of RNA biology – it can tell you which RNA sequences contain a biochemically modified residue, for example. However, longer reads (for instance, using sequencing technologies offered by Oxford Nanopore and Pacific Biosciences) can now help to determine how common a particular modification is in the cell, and whether changes in one part of an RNA molecule correlate with changes in another.

Light-up aptamers are single-stranded DNA or RNA molecules that were developed in the lab to bind to fluorescent dyes. They are RNA analogues of the green fluorescent protein that is produced in some marine animals, and when these aptamers bind to the dyes, their fluorescence intensity increases. This enables researchers to track, for example, the formation of intracellular RNA clusters that contribute to neurodegenerative diseases.

Earlier light-up aptamers were unreliable: their signals were dim, and sometimes the aptamers didn't work at all because the sequences misfolded when fused with the target RNA. But several groups have developed new types of fluorescent RNA, and in papers and talks I've seen a huge push to improve the brightness of existing aptamers and create variants that glow in different colours.

My lab has used chemical footprinting methods to study RNA folding in the cell. Many disorders are associated with changes in RNA structure, but that has been really hard to tease apart. Now we are turning to long-read sequencing and light-up aptamers to study RNA–protein aggregates in diseases including cancer, metabolic syndromes and Alzheimer's. Using these technologies, we can better correlate cell death and other disease features with what's happening to RNA molecules in the cell.

**Sarah Woodson** is a biophysicist at Johns Hopkins University in Baltimore, Maryland.

### ELHANAN BORENSTEIN DECODING THE MICROBIOME

Over the past decade, methods for sequencing the genetic content of microbial communities have probed the composition of the human microbiome. More recently, scientists have tried to learn what the microbiome is doing by integrating information about genes, transcripts, proteins and metabolites. Metabolites are especially interesting: they could offer the closest understanding of how the microbiome affects our health, because many host–microbiome interactions occur through the metabolites that bacteria generate and consume.

There has been an explosion of microbiome–metabolome studies looking at, for instance, a set of stool samples – identifying the species present in each sample and their abundances through metagenomic sequencing, and using mass spectrometry and other

**“Metabolites could offer the closest understanding of how the microbiome affects our health.”**

technologies to measure the concentrations of different metabolites. By combining these two profiles, the hope is to understand which member of the microbiome is doing what, and thus whether specific microbes determine the level of certain metabolites.

But these data are complex and multi-dimensional, and there might be a whole web of interactions, involving multiple species and pathways, which ultimately produce a set of metabolites. Scientists have published computational methods to link microbiome and metabolome data and to learn these quirks and patterns. Such methods range from simple correlation-based analyses to complex machine-learning approaches that use existing microbiome–metabolome data sets to predict the metabolome in new microbial communities, or to recover microbe–metabolite relationships.

Our lab takes a different strategy. Rather than apply statistical methods to find microbe–metabolite associations, we build mechanistic models of how we think a specific microbial composition affects the metabolome, and use these as part of the analyses themselves. In effect, we are asking: on the basis of genomic and metabolic information, what do we know about each microbe's ability to produce or take up specific metabolites? We can then predict the potential of a given collection of microbes to produce or degrade specific metabolites, and compare those predictions with actual metabolomic

data. We showed that this approach avoids the pitfalls of simple correlation-based analyses<sup>4</sup>, and will release a new version of the analysis framework in the coming months.

Such studies could improve microbiome-based therapies by identifying, for example, specific microbes responsible for producing too much of a harmful metabolite or too little of a beneficial one.

**Elhanan Borenstein** is a computational systems biologist at Tel Aviv University, Israel.

### CHRISTINA CURTIS COMPUTING CANCER

When it comes to cancer, we cannot see the process by which the disease forms, only its end point: we sample a tumour when it has become clinically detectable. By then, the tumour has acquired many mutations, and we're left to work out what happened.

Our team built a computational model to explore the dynamics of tumour progression while accounting for tissue spatial structure. With this model, you can simulate a range of scenarios and generate ‘virtual tumours’ with patterns of mutation that mimic patient data. By comparing simulated data with actual genomic data, it's possible to infer which parameters probably gave rise to a patient's tumour.

I'm excited about complementing these inferential approaches with direct measurements of tumour lineage and phenotype using emerging barcoding and recording methods. Advances in the past two years include evolving

CRISPR-based barcodes that can record the fate of cells during mammalian development<sup>5,6</sup>. Other techniques use image-based detection of DNA barcodes through *in situ* expression of RNA, thereby capturing cellular lineage, spatial proximity and phenotypes<sup>7</sup>.

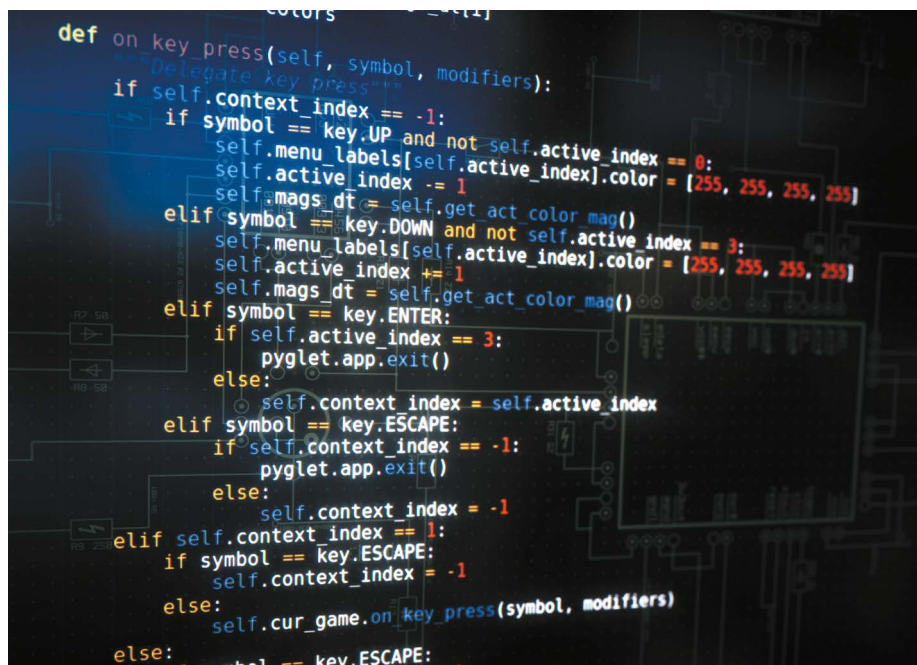
In a study that modelled the growth of tumours in colon cancer<sup>8</sup>, we used tumour sequence data and simulations to study relationships between primary and metastatic tumours. These inferential analyses indicated that the vast majority of cancers had spread when the primary tumour comprised barely 100,000 cells – too small to detect using standard diagnostic methods such as colonoscopy.

With better sensitivity and scalability, a blend of modelling and measurement methods could track both lineage and spatial relationships during tumour formation, giving insight into cancer's origins, including how specific mutations influence cellular fitness and fuel the disease's progression.

**Christina Curtis** is a computational and systems biologist at Stanford University, California.

### ALEX NORD ENHANCING GENE THERAPY

We're now about 15 years into large-scale experiments to map enhancers and other regulatory DNA sequences that control how genes are read out by cells and organs. Although more work is needed to complete these maps, we're at the point at which we can harness our understanding to control the genome more precisely.



Software code can be used to build models that simulate tumour development.



At the Society for Neuroscience annual meeting last October in Chicago, Illinois, I co-chaired a session that focused on identifying enhancer sequences and using them to control gene expression in specific cell types in the brain. One approach delivers engineered viruses into the brain to test thousands of enhancers for the gene-expression profile of interest. In 2019, researchers at the Allen Institute for Brain Science in Seattle, Washington, used this strategy to look for enhancers in specific cortical layers in the human brain<sup>9</sup>. And a team from Harvard University in Cambridge, Massachusetts, used an RNA-sequencing-based method to find enhancers that act only in specific interneurons, a type of nerve cell that creates circuits<sup>10</sup>.

Once enhancer sequences are identified, scientists can use them to drive expression in particular cell types for gene-therapy applications. In disorders caused by the inactivation or deletion of one copy of a gene, CRISPR–Cas9 gene-editing tools can target transcriptional activators to the gene's enhancer to turn up expression of the working copy. Research in mice suggests these approaches can correct gene-expression deficiencies that lead to obesity and to conditions such as fragile-X, Rett and Dravet syndromes<sup>11</sup> – the latter a severe form of epilepsy that my lab is working on. In the coming year, I think we'll still be curing mice, but there is a lot of industry investment in this technology. The hope is that we can use these methods to transform how gene therapy is done in humans.

**Alex Nord** is a geneticist at the University of California, Davis.

## J. CHRISTOPHER LOVE SINGLE-CELL SEQUENCING

I'm interested in how we bring medicines to patients faster and more accessibly. The technologies required are multifaceted. On the one hand, there's discovery – for example, single-cell sequencing methods. On the other hand, there's the matter of getting the technology to the patient – the manufacturing part. This is particularly relevant to medicines for rare diseases or for small populations, and is even applicable to global access to medicines we already have.

On the discovery front, we've worked with colleagues at the Massachusetts Institute of Technology (MIT) in Cambridge to develop a portable, inexpensive platform for high-throughput, single-cell RNA sequencing<sup>12</sup>. But it's still challenging to get sufficient resolution to distinguish between immune-cell subtypes, for instance, with different roles and antigen specificities. Over the past year



**Activated T cells from human blood.**

or so, we've enhanced single-cell genomic sequencing in several ways. First, we came up with a method for detecting low-expression transcripts more efficiently<sup>13</sup>. And for T lymphocytes specifically, we designed a protocol that links each cell's gene-expression profile with the sequence of its unique antigen receptor<sup>14</sup>.

Meanwhile, a team at the Dana Farber Cancer Institute in Boston, Massachusetts, has published a clever library-screening strategy to address the other side of the equation – working out which antigen a particular T-cell receptor recognizes<sup>15</sup>.

With MIT collaborator Alex Shalek and others, I have started a company, Honeycomb Biotechnologies, to commercialize our single-cell RNA-sequencing platform. Instead of having to spin down cells in a centrifuge, stick them in a tube, freeze it in liquid nitrogen and ship it from Africa, say, you could just ship an array of single-cell-sized wells – something the size of a USB thumb drive. That could make single-cell storage and genomic profiling possible for just about any sample anywhere in the world.

**J. Christopher Love** is a chemical engineer at the Koch Institute for Integrative Cancer Research at MIT in Cambridge, Massachusetts.

## JENNIFER PHILLIPS-CREMINS LINKING GENOME STRUCTURE AND FUNCTION

When you stretch out a single cell's DNA end to end, it's roughly 2 metres long – yet it has to fit into a nucleus with a diameter smaller than the head of a pin. The folding patterns cannot be random; chromosomes form 3D structures that must be spatially and temporally regulated across an organism's lifespan.

With genomics and imaging advances over the past decade, we can now create ultra-high-resolution maps of how the genome folds. Now the big question is,

what is the function of each of these folding patterns? How do they control fundamental processes such as gene expression, DNA replication and DNA repair?

Several synthetic-biology approaches could allow us to fold and probe the genome across a range of length- and timescales. One method, CRISPR-GO, can carry pieces of DNA to specific compartments on or in the nucleus<sup>16</sup>. This will allow scientists to ask how the nuclear placement of DNA sequences governs gene function.

Another is our lab's light-activated dynamic looping (LADL) tool, which uses light and CRISPR–Cas9 to tether specific pieces of DNA together on demand over long distances<sup>17</sup>. This can bring an enhancer into direct contact with a target gene thousands or even millions of bases away, so we can directly assess that regulatory sequence's function: does expression of its target gene go up or down, and to what degree? The technology allows precise spatio-temporal control over gene expression, which is critically disrupted in many diseases.

A third system, CasDrop, uses another light-activated CRISPR–Cas9 system to pull specific pieces of DNA into subnuclear membraneless 'condensates'<sup>18</sup>. Their function in cells has been hotly debated since they were discovered a few years ago.

What inspires me for the future is that we can couple these 3D genome-engineering tools with CRISPR-based live-cell imaging approaches, so that we can both engineer and observe the genome in real time in cells.

Function could drive structure. Or structure could drive function. This is a great mystery that these engineering tools will allow us to answer.

**Jennifer Phillips-Cremins** is an epigeneticist and bioengineer at the University of Pennsylvania, Philadelphia.

1. Liu, N. et al. *J. Am. Chem. Soc.* **141**, 4016–4025 (2019).
2. Wei, H. et al. *J. Struct. Biol.* **202**, 170–174 (2018).
3. Schaffer, M. et al. *Nature Meth.* **16**, 757–762 (2019).
4. Noecker, C. et al. *mSystems* <https://doi.org/10.1128/mSystems.00579-19> (2019).
5. Chan, M. M. et al. *Nature* **570**, 77–82 (2019).
6. Kalhor, R. et al. *Science* **361**, eaat9804 (2018).
7. Askary, A. et al. *Nature Biotechnol.* **38**, 66–75 (2020).
8. Hu, Z. et al. *Nature Genet.* **51**, 1113–1122 (2019).
9. Mich, J. et al. Preprint at bioRxiv <https://doi.org/10.1101/555318> (2019).
10. Hrvatin, S. et al. *eLife* **8**, e48089 (2019).
11. Colasante, G. et al. *Mol. Ther.* <https://doi.org/10.1016/j.ymthe.2019.08.018> (2019).
12. Gierahn, T. M. et al. *Nature Meth.* **14**, 395–398 (2017).
13. Hughes, T. K. et al. Preprint at bioRxiv <https://doi.org/10.1101/689273> (2019).
14. Tu, A. A. et al. *Nature Immunol.* **20**, 1692–1699 (2019).
15. Kula, T. et al. *Cell* **178**, 1016–1028 (2019).
16. Wang, H. et al. *Cell* **175**, 1405–1417 (2018).
17. Kim, J. H. et al. *Nature Meth.* **16**, 633–639 (2019).
18. Chin, Y. et al. *Cell* **175**, 1481–1491 (2018).

### Interviews by Esther Landhuis.

These interviews have been edited for length and clarity.



### Where I work Erica McAlister

Photographed for *Nature* by  
Leonora Saunders.

**W**e don't know the exact number of dead insects in the entomology collection at the Natural History Museum in London, but it's more than 34 million.

Our collections, for me, are a place of wonder. The specimens they contain are the biological heritage of the planet: splendid to look at and packed with genetic information about the past. Some have come to us from across the globe, and make me feel how small I am, as part of our biosphere.

The insect collection stretches back hundreds of years. For example, we have a robber fly caught in 1680 by the queen's gardener at Hampton Court Palace, near London.

Flies are my focus. Not only are they amazingly diverse, but they're cute. We've got stalk-eyed flies; flies that are less than a millimetre in size; and my favourites, *Mallophora* robber flies, which look like massive bumblebees and are highly venomous. I also have a soft spot for botflies, one species of which (*Cephalopina titillator*) matures in camels' nostrils.

The collection isn't static; there's so much research going on. We're always updating

nomenclature, revising evolutionary family trees and describing new species.

The museum lends specimens by post, and we host not just scientists, but visitors such as designers looking for inspiration. We're also trying to digitize the entire collection so that anyone can access it.

I'm collaborating with Mara Lawniczak at the Wellcome Sanger Institute in Hinxton, UK, on what we call Project Neandersquito. We're trying to recover genomes from mosquito samples collected over the past century. In the past, people would cut off legs or destroy whole specimens – which fills a curator like me with terror. Instead, we are washing the specimens with chemical solutions to extract DNA.

Genetic analysis will help us to distinguish between old mosquito specimens that look similar, and to learn how populations have changed. For example, we hope to see when genes for insecticide resistance arose.

**Erica McAlister** is a senior curator at the Natural History Museum in London, UK, and author of *The Secret Life of Flies* (Firefly Books, 2017). **Interview by Amber Dance.**